

Bayesian Inference for Comparative Research Author(s): Bruce Western and Simon Jackman Source: The American Political Science Review, Vol. 88, No. 2 (Jun., 1994), pp. 412-423 Published by: American Political Science Association Stable URL: <u>http://www.jstor.org/stable/2944713</u> Accessed: 13/06/2011 15:41

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=apsa.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Political Science Association is collaborating with JSTOR to digitize, preserve and extend access to The American Political Science Review.

BAYESIAN INFERENCE FOR COMPARATIVE RESEARCH

BRUCE WESTERN Princeton University SIMON JACKMAN University of Rochester

Regression analysis in comparative research suffers from two distinct problems of statistical inference. First, because the data constitute all the available observations from a population, conventional inference based on the long-run behavior of a repeatable data mechanism is not appropriate. Second, the small and collinear data sets of comparative research yield imprecise estimates of the effects of explanatory variables. We describe a Bayesian approach to statistical inference that provides a unified solution to these two problems. This approach is illustrated in a comparative analysis of unionization.

widely used tool in quantitative comparative politics is the linear regression model, estimating relationships between institutional and economic variables observed for a set of countries constituting some population of interest. A brief survey of recent comparative research finds no shortage of work of this type. For instance, Remmer (1991) identifies economic crisis as a source of electoral volatility in Latin American countries in the 1980s; Robertson (1990) presents evidence that institutions moderating the transaction costs of negotiations between capital and labor reduce industrial conflict; and Golden (1993) finds that union wage moderation in industrialized democracies results from coordinated collective bargaining.

Although the methodological advantages are well documented (Jackman 1985), we identify two distinct problems of statistical inference endemic to this type of quantitative comparative research. First, because the data constitute all the available observations from a population, conventional inference based on the long-run behavior of some repeatable data mechanism is not appropriate. Second, because data sets tend to be small and collinear, the effects of explanatory variables are estimated imprecisely. In most applications the first problem is generally ignored and the second is solved by imposing restrictions on the regression model to obtain "sensible" results.

The Bayesian model of statistical inference provides a unified solution to these two distinct problems of quantitative comparative research. Two features of Bayesian inference are important in this context. First, probability is conceived subjectively as characterizing a researcher's uncertainty about the parameters of a statistical model. This subjective probability concept seems particularly useful in comparative settings, where the data are typically convenience samples not generated by a known probability mechanism such as random sampling. Second, Bayesian inference allows the inclusion of other information, in addition to the quantitative sample information. Again, this seems useful in the comparative context where rich historical material is commonly available, sparking ideas for researchers, but formally discarded in the final analysis. In the Bayesian approach, this historical material can be formally incorporated into the analysis.

We begin by detailing why the nonstochastic and weak data typical of comparative research are problematic for conventional linear regression analysis. We next introduce some ideas about Bayesian inference and describe Bayesian regression analysis. These ideas are applied to a comparative analysis of unionization—the subject of a recent exchange in the *American Political Science Review* between Michael Wallerstein and John Stephens (Stephens and Wallerstein 1991; Wallerstein 1989). The Bayesian regressions we present are supplemented with a sensitivity analysis that investigates how our conclusions depend upon the sample data and our prior beliefs.

Although we survey some ideas in Bayesian statistics, a full introduction to the Bayesian approach to statistical inference is beyond our scope here. (Good book-length treatments are Pollard 1986, Lee 1989 and, at a slightly higher level, Press 1989.) Our intention here is to spotlight problematic but ignored issues of statistical inference in a common area of application. The Bayesian alternative that we present represents just one way forward to address these inferential problems.

TWO FEATURES OF COMPARATIVE DATA ARE PROBLEMATIC

A common design in comparative research generates data with two important characteristics: (1) the data constitute all the available observations from a population; and (2) the data, because of small sample size and collinearity, tend not to be very informative about the statistical parameters being estimated. Although our focus on comparative research is stimulated by the preponderance of studies with data of this type, our arguments are generalizable to other areas using data with these characteristics. For example, studies comparing the American states could well be another area of application (e.g., Barrilleaux and Miller 1988; Erikson, McIver, and Wright 1987).

Nonstochastic Data

Unlike analysts of survey research data or experimenters who randomly assign subjects to treatments, comparativists often collect all the available observations from some population of interest. For example, populations including "advanced industrial soci-eties" (Wallerstein 1989), "contemporary democratic political systems" (Powell 1982, 2), and the "affluent market-oriented democracies" (Swank 1988, 1121) have all been studied by comparative researchers. In contrast to experimental or social survey data, comparative data are not generated by a repeatable data mechanism. Repeated applications of survey or experimental designs, on the other hand, yield new data sets with more information about the process under study. In comparative research, once all the data are collected from a population, further applications of the data collection process do not yield more information. Data of this type, generated by a nonrepeatable and unknown probability process, are described by Freedman and Lane (1983) as "nonstochastic." We adopt their terminology here.

To see why nonstochastic data create problems for conventional statistical inference, we briefly review the concept of frequentist probability. This idea is the workhorse of conventional, or frequentist, statistics. In 1837, Denis Poisson defined probability as the limiting distribution of a long-run relative frequency. If, for example, a coin is tossed *n* (a very large number) times and shows *m* heads, we can write the prob(heads) $\equiv p = m/n$, assuming that m/n coverges. In this case, the probability of heads, *p*, is the proportion of heads that will be observed when *n* grows to infinity. The probability, *p*, like size or weight, describes an objective characteristic of the coin (Barnett 1982, chap. 3; Leamer 1978, chap. 2).

Frequentist statistical inference assumes that data are generated by a repeatable mechanism such as the coin flip. Von Mises provides the paradigmatic state-ment of this view: "In order to apply the theory of probability we must have a practically unlimited sequence of uniform observations" (quoted in Barnett 1982, 76). A sample observation is thus just one possible result from many possible draws from a probability distribution. In practice, these draws are implemented by probability sampling (as in a social survey) or random assignment (in an experiment). Frequentist inference makes conclusions about a parameter (perhaps a mean or a regression coefficient) obtained from a sampling or assignment process, as if that process was repeated a large number of times. For example, a random sample survey of American adults may indicate that mean income in the United States is \$35,000. Assuming (rather implausibly) that income is normally distributed, we could estimate a 90% confidence interval for our sample mean, perhaps [\$15,000, \$55,000] for a modestly sized sample. Using conventional frequentist inference we can conclude that intervals like the one calculated would cover the true (population) mean income 90% of the time for repeated applications of the sampling procedure. The "repeated sampling" inference tells us neither whether the population mean lies within the estimated interval nor even with what probability the mean lies in the interval. Specifically, the frequentist inference does not entitle us to claim there is a 90% chance that the true mean falls within the estimated interval. The only available conclusion concerns the long-run behavior of a statistic—in this case, the 90% confidence interval.

While frequentist inferences from survey data allow conclusions about probability sampling processes, nonstochastic data create problems for frequentist inference because they are not generated by a repeatable mechanism. Frequentist inference is simply unrealistic given the manner of data collection. Comparative researchers sometimes show an uneasy awareness of this problem. In their analysis of advanced industrial democracies, Lange and Garrett frankly report that they "adhere to traditional standards [i.e., significance tests] while remaining unsure of their applicability" (1987, 268). In a similar vein, Weede endorses the use of *t*-statistics for testing hypotheses in his study of welfare state size in industrial democracies, not only for methodological reasons but also because "it is almost universal practice in econometrics and the social sciences" (1986, 518). More confusion is revealed in Gorin's comparative study of inequality that "cannot utilize probability theory to ascertain the level of significance" for want of "random samples"; he goes on to note that asterisks indicate a ".05 level of significance" (1980, 3:153).

Comparative researchers' discomfort with frequentist inference is well founded because *frequentist inference is inapplicable to the nonstochastic setting.* It is simply not relevant for the problem at hand to think of observations as draws from a random process when further realizations are impossible in practice and lack meaning even as abstract propositions. In short, frequentist inference answers a question that comparative researchers are not typically asking (see also Freedman and Lane 1983, 189).

Two objections might be raised at this point. One challenge might be that statistical inference is unnecessary in the nonstochastic setting because all the available information is collected. This position is certainly valid, but it still commits the researcher to a substantive theory of how the data were generated. In particular, if there is no uncertainty associated with the data and if statistical inference is expendable, the researcher is effectively claiming that things could not have been different, that the data were generated by a completely deterministic process. This suggests a thought experiment: If we could set in motion, once again, the historical conditions that gave rise to the data, would the ensuing process generate a data set identical to the one actually obtained? Commitment to this type of determinism is the cost of abandoning statistical inference (Berk, Western, and Weiss 1993).

The second objection holds that although the data mechanism is not repeatable, it can be treated as if it

were. In this view, the social world randomly draws observations from a set of all possible observations, or a "superpopulation" (Cochran 1953, 169). The data are one realization of all possible data sets that might have been collected. Or as one reviewer put it, "each country's history is one draw of a distribution of possible histories." While this assumption avoids a deterministic theory of the data process, it is highly speculative compared to positive knowledge about a sampling procedure or, in Fisher's uncompromising phrase, "the physical act of randomization" (quoted in Freedman and Lane 1983, 197). Even more troubling, however, is the conclusion that conventional inference allows in this instance. Take the example of a confidence interval for a mean where we can conclude that under repeated realizations (which are acknowledged to be impossible), the interval would cover the true mean 90% of the time. We have no way of knowing whether the current interval is one of the fortunate 90% and no possibility of further replications. (Our focus here has been on the confidence interval because this is the inferential tool on which we shall rely in our application and the one that implicitly underlies classical hypothesis tests. Similar criticisms of classical hypothesis testing and other aspects of frequentists inference are reviewed in Barnett 1982, 180-88).

Weak Data

The problem of statistical inference in comparative research is exacerbated by the weakness of comparative data. These data are weak in the sense that they provide little information about parameters of statistical models. The uninformative character of the data is reflected in large standard errors and consequently large *p*-values.

Weak comparative data have two distinct causes. First, the sample sizes characteristic of comparative research are small in relation to the number of parameters being estimated. While comparativists frequently make do with less than a hundred data points, survey researchers frequently have more than a thousand. For example, comparative researchers have defined the population of advanced capitalist democracies to include between 15 and 21 countries (Lange and Garrett 1985; Wallerstein 1989). Sometimes, even where samples sizes are fairly large, a relatively large number of parameters must be estimated (e.g., Williams 1991). Ideally, if more data could be collected this would be one solution to the uninformative data problem. As the sample size increases in relation to the number of parameters, the estimated variances of the regression coefficients would tend to decline.

Collinearity is the second cause of weak comparative data. In a comparative setting, collinearity is common because the explanatory variables are themselves often causally related. When explanatory variables are collinear, they carry little independent information about the various regression coefficients. A regression on collinear predictors can yield estimates with unexpected signs and large standard errors.

To foreshadow some Bayesian ideas, collinearity provides no statistical difficulties except in very extreme instances (of exact or near-exact linear dependencies) which are seldom encountered in practice. The least squares estimators retain their properties, and statistical inference can proceed as normal. Why, then, is collinearity regarded as problematic? The difficulties caused by collinearity—large standard errors and unexpected signs for regression coefficients-are only problems with respect to prior expectations about the signs and variances of the coefficients (Learner 1978, 170). As Learner argues, collinearity is thus not a statistical problem but a problem of the interpretation of multidimensional evidence (pp. 170–73). As a tool of interpretation, the least squares estimator in the presence of collinearity does not allow us to distinguish information about one coefficient from information about another. Instead, the least squares estimator is informative about linear combinations of coefficients. The introduction of nonsample information allows sample information to be allocated among the coefficients according to substantive criteria.

A more familiar approach to collinearity involves constraining coefficients to zero (excluding them from the analysis). This is an informal way of introducing information to obtain sharper estimates of other coefficients. The researcher is implicitly saying, "To obtain a sensible estimate for the effect of x_1 , I must estimate the model on the basis of the substantive claim that x_2 had no effect at all." The substantive claim that x_2 has no effect allows the weight of the sample evidence to be allocated in favor of x_1 . Here, of course, a "sensible" estimate of the regression coefficient for x_1 is based on (typically unstated) prior expectations. Dropping x_2 introduces known specification errors, and in part defeats the motivation for statistical control that led to multivariate techniques in the first place. (The papers in Granger 1990 on model specification provide an elaborate treatment of this issue from Bayesian and non-Bayesian perspectives in econometrics.)

In sum, the weak evidence of comparative research that results from small data sets and collinearity yields weak or fragile inferences that are highly sensitive to the model specification. This is the natural outcome of an analysis based on little evidence. The only solution to this problem is the introduction of more information. But as a practical matter, more information is frequently not available. Data have been collected from all the countries in the sample. The growing popularity of pooled cross-sectional times series designs suggest one remedy to the weak data problem (e.g., Alvarez, Lange, and Garrett 1991; Beck et al. 1993; Radcliff 1992; Swank 1992); but additional time points often provide little new information because the processes being studied show far more cross-national than longitudinal variation (e.g., Wallerstein 1989, 482). Institutions, for example, do not change much over time but show considerable

cross-national variation. Furthermore, the collinearity problem is distinct from small sample size and can afflict the pooled cross-sectional design irrespective of the number of observations. However, while extra quantitative information is typically unavailable, large and substantively rich stores of qualitative information from comparative and historical studies are often present but not available in a form suitable for analysis. Bayesian procedures enable the weak quantitative information of comparative research to be pooled with the qualitative information to obtain sharper estimates of regression coefficients. Although we shall focus on situations in which the data are weak, Bayesian methods can also be applied (generally with less consequence) where the data are highly informative.

BAYESIAN INFERENCE

The Bayesian approach to statistical inference involves pooling nonsample (or "prior") information with sample data to formulate posterior subjective probability statements about the parameters of a statistical model. This description of Bayesian inference signals two aspects of the Bayesian approach that distinguish it from conventional inference. First, Bayesian inference is built upon a subjective probability concept. Second, Bayesian inference allows the introduction of prior information in addition to the sample information to make statistical inferences.

Subjective Probability

In contrast to the frequentist concept that refers to objective probability features of the world, subjective probability refers to a person's *degree of belief* in an uncertain event. Subjective probability is thus a personal statement of certainty or confidence, rather than a fact characterizing an object in the external world. Subjective probabilities share the same axioms as objective probabilities and are consistent with formal rules for rational decision making (Barnett 1982, chap. 3). The difference between subjective and objective probability resides in whether the probability statement refers to an individual's personal feeling of confidence, or whether it characterizes an objective feature of the world.

Prior Information in Regression

In Bayesian inference, researchers' subjective probability assessments of the parameters of a statistical model are pooled with the sample data to arrive at posterior probability statements about those parameters. The prior information is represented in a probability distribution. In the normal regression case that we are concerned with, this prior probability distribution can be summarized by prior means and variances for the regression coefficients. (Tanner 1993 considers more complicated applications.) The posterior probability statements express the researchers' degree of belief in the parameters, given the data and the prior, replacing conventional inference about the distribution of coefficient estimates in repeated sampling. We can begin to see how prior information is incorporated into a data analysis by first reviewing how two data sets might be pooled in a regression analysis.

Two data sets, (X_1, y_1) and (X_2, y_2) , can be pooled for a regression analysis as follows:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Following Learner, if we write σ_1^2 and σ_2^2 for the two residual variances, the least squares estimate of β from the pooled data is

$$b_{pooled} = (\sigma_1^{-2} X_1' X_1 + \sigma_2^{-2} X_2' X_2)^{-1} (\sigma_1^{-2} X_1' X_1 b_1 + \sigma_2^{-2} X_2' X_2 b_2), \quad (1)$$

where $b_1 = (X'_1X_1)^{-1}X'_1y_1$, and $b_2 = (X'_2X_2)^{-1}X'_2y_2$ (1978, 76).

Although equation 1 may look a bit cumbersome, the pooled estimate for β is just a (matrix) weighted average of the two sets of least squares estimates obtained from each of the data sets.

Equation 1 is useful from the Bayesian point of view, because it suggests how sample and prior information can be combined to obtain posterior distributions for regression coefficients. The regression coefficients estimated from the first data set, b_1 , can be replaced by a set of regression coefficients that the researcher believes a priori are most probable. These are the prior means, written b^* , of the prior probability distributions we have referred to. Note that the inverse of the first part of the first term in equation 1 is the estimated variance of the coefficients from the first data set. This variance can be substituted by the subjective Bayesian prior variance, V^* , for the coefficients. Uncertainty about the prior means is reflected in the specification of V^* . Larger variances for a prior mean imply greater uncertainty. Prior covariances-off-diagonal terms in the prior covariance matrix-can also be specified if beliefs about one coefficient depend on beliefs about others —say, if one is estimating main and interactive effects in a regression equation and if confidence in the interaction effects depends on confidence in the main effects (e.g., Lange and Garrett 1985). The prior covariances, expressing dependent beliefs about the regression parameters, thus do not directly relate to prior beliefs about relationships in the sample data. Because beliefs about the coefficients are generally independent, the prior covariance matrix for the regression coefficients is often specified to be diagonal.

Substituting the prior information into equation 1 provides an expression for the posterior mean based on a combination of the prior and sample information:

$$b = (V^{*-1} + \sigma^{-2}X'X)^{-1}(V^{*-1}b^* + \sigma^{-2}X'y)$$
 (2)

where *X* and *y* designate the sample information (the subscripts are now redundant) and the residual variance, σ^2 , is estimated from the sample data. (Note that when the expression for the sample parameter estimate is expanded in equation 1, the cross-product matrices in the second term cancel out, leaving X'y.) The variance of *b* is given by the first term of equation 2 and standard deviations for the posterior coefficients are the square roots of the diagonal elements of this matrix. An alternative route to these expressions for the posterior means and variances of the regression coefficients is to multiply a normal prior distribution by a normal likelihood for the sample data, treating the residual variance as known. This is an application of Bayes Theorem, and it yields a multivariate normal posterior distribution for the coefficients. The posterior means and variances of this multivariate distribution are identical to the expression we obtained in equation 2. The pooling exposition highlights, however, how subjective prior information can supplement the weak sample data common in comparative research.

Approaching the Bayesian regression problem in this way, the posterior is simply a pooled estimate based on the sample data and the "data" represented by the researcher's prior beliefs. Bringing prior information to bear on an inferential problem via Bayes Theorem involves the same procedure as if one were to collect more data: in both cases, the researcher pools the sample information with the prior distribution or additional data using equation 1 or equation 2.

The resolution of collinearity problems through the introduction of prior information is made clear by the expression for the posterior variance—the first term of equation 2. If prior means are set to zero, a diagonal prior covariance matrix will improve the conditioning of the explanatory variables, resulting in regression coefficients with smaller variances. This estimator is a special case of the generalized ridge estimator (Belsley 1991, 300), except that in the Bayesian case the ridge for the sample cross-products matrix is chosen according to substantive criteria rather than arbitrarily. Bayesian regression thus solves the problem of collinearity by using more information than ordinary least squares, generating smaller standard errors for the regression coefficients.

A convincing Bayesian analysis goes further than just calculating the posterior distributions. Because the two ingredients for Bayesian regression—the prior and the sample information—influence the final results, a thorough analysis should explore the sensitivity of these results to the prior and the data. Robust findings yield substantive conclusions that are insensitive to small changes in the sample and nonsample information. These ideas will be detailed below.

In sum, Bayesian regression analysis provides answers to the problems of comparative research that we have identified. First, subjective probability replaces the conventional concept, providing a more realistic basis for analyzing data collected without a repeatable data mechanism. Second, more information is brought to the analysis through the prior distribution. Weak data resulting from small samples and collinearity (common in comparative research) are bolstered by the prior, increasing the precision of the estimated regression coefficients. Finally, the robustness of Bayesian results can be explored through an analysis of the sensitivity of the posterior to the sample data and the prior information.

BAYESIAN REGRESSION ANALYSIS IN A MODEL OF UNION DENSITY

We illustrate Bayesian regression analysis in the context of a recent exchange between Wallerstein and John Stephens. Wallerstein and Stephens are concerned to explain cross-national variation in union density-the percentage of a work force that are union members. Wallerstein argues that the size of the civilian labor force is a key determinant of union density. Stephens claims that density depends on industrial concentration. These two predictors correlate at -.92 and Stephens writes that "because of multicollinearity, economic concentration and the size of the labor force could not be entered in the same equation" (Stephens and Wallerstein 1991, 945-46). Of course, as noted, no statistical argument prevents the inclusion of both predictors in a regression on union density. Stephens is simply observing that because of the correlation between labor force size and industrial concentration, the data will not be very informative about coefficients for either variable, and the results may be sharply at odds with prior expectations. The problem of collinearity results partly from incomplete data for economic concentration. Stephens uses information about logged gross domestic product to impute missing observations for 9 of the 20 countries in the sample, yet gross domestic product also correlates highly with labor-force size. For this analysis, we bracket the issue of missing data, although it should be noted, with Wallerstein, that Stephens's imputation strategy provides only weak information about economic concentration (pp. 949-50).

Two other features of the Stephens-Wallerstein data are also noteworthy. First, the sample is not generated by any known probability mechanism such as random sampling. The 20 countries comprising the sample were chosen because they experienced a continuous history of political democracy since World War II (Wallerstein 1989, 489). In the absence of some rather fanciful assumptions about probability features of the data-generating mechanism, the classical model of statistical inference is inappropriate. Second, the weak data problem that results from collinearity is compounded by the small sample that provides little statistical power. Stephens implicitly recognizes this problem in the analysis by using a relatively large level of statistical significance (.1). Here, then, is a classic example of weak data generated by an unknown probability mechanism in a

Prior Means (Standard Deviations) and Their	· Substantive Interpretations	for Bayesian Regression Analysis of
Union Density	_	

VARIABLE	PRIOR	SUBSTANTIVE INTERPRETATION			
Wallerstein's priors					
Left government	.3 (.15)	One year of left-wing government increases union density by about 1 percentage point. A year of left-wing government may increase union density by as much as 2 percentage points of union density, but its effect is almost certainly not negative.			
Logged labor-force size	-5 (2.5)	Doubling the size of the labor force would reduce union density about $ln(2) \times 5 \approx 3.5$ percentage points. This increase in labor-force size may generate a union decline as big as 7 percentage points, but a growing labor force is unlikely to increase union density.			
Economic concentration	0 (10 ⁶)	The diffuse prior indicates that the researcher has no strong prior beliefs about the sign or magnitude of an effect. When the explanatory variables are uncorrelated, the diffuse prior yields posteriors that are approximately given by the sample data.			
Stephens's priors					
Left government	.3 (.15)	Like Wallerstein's prior, one year of left-wing government increases union density by 1 percentage point.			
Logged labor-force size	0 (10 ⁶)	Diffuse prior.			
Economic concentration 10 (5)		If economic concentration were to increase by 100% in relation to the United States, union density would increase by 10 percentage points. This increase in the concentration ratio may generate a density increase as large as 20 percentage points, but any increase in concentration is unlikely to decrease union density.			

Note: Left government is measured by Wilensky's (1981) cumulative index of left-wing government; logged labor-force size is the natural log of the size (in thousands) of the dependent labor force in the year that union density is measured; and economic concentration is measured by the four-firm concentration ratio, in proportion to the United States.

comparative setting. Our data are taken from, and described by, Stephens (Stephens and Wallerstein 1991). Note that labor-force size is measured on the log scale, so that the coefficients express the changes in union density in response to percentage changes in labor-force size. Economic concentration is measured as a ratio to U.S. economic concentration, so that the coefficients express the change in union density as a result of a country's change in economic concentration in relation to the United States.

What light can Bayesian regression analysis shed on the Stephens–Wallerstein debate? From the Bayesian perspective, the source of the controversy between Stephens and Wallerstein is the data that are not strong enough to allow a convincing resolution one way or another. Both researchers informally enlist more information to support their arguments. In Wallerstein's case, this information is a formal theory about the relationship between labor-force size and union density. For Stephens, supplementary information comes from historical material and, in particular, the research of Kjellberg (1983). The Bayesian approach allows the information informally introduced into the analysis by Stephens and Wallerstein to enter formally through a prior distribution.

To begin this analysis, we specify two sets of prior information that represent the opinions of the two researchers. Unfortunately, neither theory is suffi-

ciently precise to unambiguously suggest a prior distribution. (We shall return to this issue.) However, it is clear that both Stephens and Wallerstein are confident that left governments assist union growth. We quantify this confidence with a prior mean of .3 and a prior standard deviation of .15 (Table 1). Because the prior 95% confidence interval— $.3 \pm (1.96)$ \times .15)—does not overlap zero, both researchers are quite confident that left government is unlikely to have a negative effect on unionization. As Wallerstein notes, this effect probably suffers from simultaneity bias because unions increase the electoral success of left-wing parties (1989, 490). Because our substantive focus is on the effects of labor-force size and economic concentration, we treat the left-government variable chiefly as a control and simply note that the coefficient will tend to overestimate the impact of left parties on union density. (The problem of simultaneity can be addressed through the prior distribution, but we reserve this issue for another paper; see, e.g., Leamer 1991.)

Prior opinions of the two researchers differ on the labor-force size and economic concentration effects. Wallerstein has a strong belief in a negative laborforce size effect, although his theory does not suggest how large this effect might be. A few quasi-experimental comparisons using information from outside the sample suggests a plausible prior mean. In 1950, Sweden and Norway shared similar union structures, ethnic homogeneity, and histories of social democracy. The Swedish labor force was about twice as large as the Norwegian, and the unionization gap was about 20 percentage points. If a fifth of the unionization gap (4 percentage points) were attributable to labor-force size, the size effect size would be $4/\ln(2) \approx 5.5$. We get a similar number if we think that about a fifth of postwar American union decline is due to the near-doubling of the U.S. labor force. Choosing a size effect to explain a fifth of the variability in union density reflects, rather arbitrarily, our belief in the quality of the model specification. Fourfifths of the variability in union density is left to the effects of other causes. Confidence in the sign of the effect is reflected in a prior standard deviation that yields a 95% confidence interval that excludes zero.

Stephens believes that economic concentration increases union density, although a prior mean is not obvious from his discussion. Again, we develop a prior based on a quasi-experimental comparison using nonsample information-this time, examining declining economic concentration and union density in the United Kingdom through the 1980s. Measuring economic concentration as the average size of British manufacturing establishments as a ratio of the size of American firms at a fixed point in time, economic concentration in Britain declined by about .3 from the late 1970s to the late 1980s. (See the United Nations Yearbook for Industrial Statistics for 1982 and 1991.) Union density in this period declined by about 15 percentage points. If economic concentration generated about a fifth, or 3 percentage points, of the decline, a plausible prior mean would be (3/.3) = 10percentage points. Confidence in this effect is again supplied through by a prior distribution whose 95% confidence interval does not overlap zero. Note that the style of reasoning behind the specification of these priors is only illustrative. More plausible or more confident specification of prior information would require a much more detailed substantive or theoretical discussion (e.g., Western 1994).

We must also specify a prior for the intercept term, Wallerstein's prior for Stephens's economic concentration variable, and Stephens's prior for Wallerstein's labor-force-size variable. Here, we introduce the idea of prior ignorance that characterizes highly uncertain beliefs about parameters. Prior ignorance about the coefficients can be defined by a zero prior mean and a very large prior variance (Leamer 1978, 62). When ignorance is defined in this way for all the coefficients of the model with a so-called diffuse prior, the posterior distribution converges on the conventional least squares result. Thus, the sample data alone will drive the analysis. If the predictors are uncorrelated and ignorance priors are placed on only a subset of the coefficients, the posteriors for those coefficients will also approach the conventional least squares estimates. In short, this ignorance prior does not so much express the belief that a regression coefficient is zero with great uncertainty: rather, the diffuse prior allows the sample information to domi-

TABLE 2

Posterior Distributions with Noninformative Prior Information in a Regression Analysis of Union Density

INDEPENDENT VARIABLES	MEAN	5TH PERCEN- TILE	95TH PERCEN- TIL E	
Intercept	97.59 (57.48)	3.04	192.14	
Left government	.27 (.08)	.15	.39	
Size	-6.46 (3.79)	-12.70	22	
Concentration	.35 (19.25)	-31.32	32.02	
Note: These results are equivalent to the ordinary least squares estimates $(N = 20)$				

nate the prior information in the calculation of the posterior distribution. We place diffuse (or "ignorance") priors on the intercept term, the economic concentration coefficient for Wallerstein's prior, and the labor-force-size coefficient for Stephens' prior.

Results from conventional least squares estimation are in Table 2. Importantly, diagnostics show that the residuals from this least squares fit are approximately normal and the effects, approximately linear (see Fox 1990, 80-87; Hastie and Tibshirani 1990, chaps. 6-7). Thus the data conform to the structure assumed in the normal likelihood function. From a Bayesian perspective these results can be interpreted as the posterior distribution for a noninformative prior distribution. In addition to the mean and standard deviation of the posterior coefficients, we also report normal approximations for the fifth and ninety-fifth percentiles from the posterior distribution (\pm 1.64 \times standard deviation for the coefficient). These percentiles are simply an alternative way of describing the dispersion and location of the posterior distribution. The results suggest a large range of plausible and negative for values for the size effect. The effect of economic concentration is less certain. The posterior mean is close to zero and the posterior distribution stretches over a large range of positive and negative values. In short, the weight of the data favor the effect of size on union density, but neither parameter is estimated with great precision.

The introduction of prior information can sharpen these estimates. Results for Wallerstein's prior are shown in the top panel of Table 3. With the informative prior, the effect of labor-force size is now estimated with greater precision. Given the prior and the data, we can be 90% certain that the coefficient for labor-force size falls between -9 and -2. The improved conditioning of the data is also reflected in the posterior distribution for the economic concentration coefficient. Although Wallerstein's prior for the concentration effect was diffuse, the posterior mean has

Posterior Distributions with Stephens's and Wallerstein's Informative Priors in a Regression Analysis of Union Density (N = 20)

INDEPENDENT VARIABLES	MEAN (S.D.)	5TH PERCEN- TILE	95TH PERCEN- TILE
Wallerstein's prior Intercept	82.43 (32.83)	28.42	136.43
Left government	.28 (.07)	.17	.39
Logged labor-force size	-5.44 (2.09)	-8.87	-2.00
Economic concentration	4.87 (12.41)	-15.54	25.28
Stephens's prior Intercept	70.82 (19.87)	38.13	103.51
Left government	.27 (.07)	.16	.38
Logged labor-force size	-4.79 (1.77)	-7.70	-1.88
Economic concentration	9.38 (4.84)	1.42	17.34

become larger and the standard deviation of the posterior has become about one-third smaller. The confidence region, however, still covers zero.

From Stephens's perspective, when prior information is allocated to the economic concentration coefficient, the posterior mean becomes large and positive and the standard deviation of the posterior distribution shrinks by about four-fifths compared to the least squares estimate. The posterior distribution is now completely in the positive range. Again, as prior information is introduced, inference about both coefficients is generally improved. Although the posterior mean for the size effect has shrunk under Stephens's prior, the confidence region is roughly the same as that obtained with Wallerstein's prior. Both sets of prior beliefs support the inference that laborforce size decreases union density. However, only Stephens's prior supports the conclusion that growing economic concentration increases union density.

Sensitivity Analysis

It might be objected here that the introduction of subjective prior information allows the researchers' prejudices into the analysis, corrupting the results that would be obtained from the sample data alone. This criticism has a long history, demonstrated by Fisher's charge that an experiment interpreted with prior information "would carry with it the serious disadvantage that it would no longer be self-contained, but would depend for its interpretation from experience previously gathered. It could no longer be expected to carry conviction to others lacking this supplementary experience" (1935, 69). A contemporary champion of Fisher's position, Bradley Efron has similarly argued that Bayesian methods "fail to reassure oneself and others that the data have been interpreted fairly" (1986, 4; emphasis original). These comments of Fisher and Efron rests on the idea that a "fair data analysis"-a data analysis without prior information-is possible. In practice, however, prior information enters most analyses through coding decisions, transformations, and unreported searches over sets of explanatory variables to obtain results that look sensible in the sense of falling within an expected range of meaningful results. While all data analysts use prior beliefs, Bayesians go some way to making these priors explicit and integrating them systematically into the analysis. In the Bayesian approach then, acknowledged subjectivity is the route to objectivity (deFinetti 1974, 1:5-6).

Still, those objecting to Bayesian inference make a powerful point. Selecting a prior is subjective in the sense that two researchers will not necessarily agree on its specification. For example, at least part of Stephens's prior belief in the effect of economic concentration is based on his historical discussion of the growth of Swedish corporatist bargaining (Stephens and Wallerstein 1991, 944). Other researchers might believe the Swedish experience is unique in the sense of being uninformative about the more general relationship between economic concentration and unionization in comparative perspective. They might argue that Stephens's prior gives undue weight to the lessons of the Swedish case. What is more, priors are specified with a degree of whimsy in applied settings (Leamer 1983). For instance, although Wallerstein's argument led us to a prior labor-force-size effect of mean 5 and standard deviation 2.5, we could readily agree to an alternative prior, say one with the same mean but standard deviation 2.7. In short, neither several researchers nor even the same researcher, in practice, will prefer one prior probability distribution to the exclusion of all others.

Because the choice of prior is subjective in this sense of attracting little consensus, it is important to investigate how the posteriors depend on the priors. If the posteriors are highly sensitive to the priors, this suggests that the sample data add little to the prior information-indeed that inferences are driven by the priors alone. A parallel argument can be made about the relationship between the sample data and the posteriors. If a small number of observations from the sample data are highly influential for the posteriors, the results are similarly unstable, reflecting information about a few cases rather than the whole of the data set in combination with the priors. In sum, because of the joint influence of the prior information and the data on the analysis, a convincing analysis investigates the sensitivity of the posterior to the priors and the data.

The sensitivity of the posterior to the prior could be examined in several ways. One sensitivity analysis,

		5TH	95TH
INDEPENDENT VARIABLES	MEAN (S.D.)	PERCEN- TILE	PERCEN- TILE
Wallerstein's prior			
Intercept	93.50 (52.02)	7.94	179.07
Left government	.27 (.07)	.15	.39
Logged labor-force size	-6.18 (3.42)	-11.81	56
Economic concentration	1.60 (17.68)	-27.49	30.68
Stephens's prior			
Intercept	80.88 (38.14)	18.15	143.62
Left government	.27 (.07)	.15	.39
Logged labor-force size	-5.42 (2.69)	-9.85	99
Economic concentration	6.11 (12.22)	-13.99	26.20

sometimes called "extreme bounds analysis," investigates variability in the posterior as prior variances are allowed to range from zero to infinity while the prior means are fixed at zero (Leamer 1983). In our application, with nonzero prior means, we simply describe the sensitivity of the posterior to a reduction in the prior information through an increase in the prior variances. An alternative approach might manipulate the prior means, although this should be pursued in the spirit of a sensitivity analysis, leaving a priori uncertainty about the regression coefficients to be expressed in their prior variances.

We report two more sets of posterior distributions for a diffuse prior with the prior variances multiplied by 10. Under the diffuse version of Wallerstein's prior, Table 4 shows that the labor-force-size coefficient becomes larger but substantially more imprecise. The 90% confidence region still excludes-but now borders on-zero. The posterior for economic concentration is now effectively centered over zero. For Stephens's diffuse prior, the posterior mean of the concentration coefficient shrinks by about a third, and the standard deviation of the coefficient increases by about three times compared to Stephens's informative prior. The 90% confidence interval now includes a large range of negative values expressing great uncertainty about the effect of economic concentration on union density. Finally, similar to Wallerstein's diffuse prior, the posterior for size gets slightly larger but substantially more imprecise com-

FIGURE 1

Sensitivity of Bayesian Regression Results to Prior Information and the Sample Data.



a country have the same influence on the final results, regardless of rior, it will fall on the solid 45-degree line. Italy is an outlier, becoming highly influential as prior information is removed from the analysis.

pared to the posteriors based on the informative priors. To summarize, inferences under both priors become considerably more uncertain as nonsample information is removed from the analysis. A strong inference about the sign of the economic concentration effect depends crucially on the informative version of Stephens's prior.

The sensitivity of the posteriors to the data can be

INDEPENDENT VARIABLES	INFORMATIVE PRIOR			DIFFUSE PRIOR		
	MEAN (S.D.)	5TH PERCENTILE	95TH PERCENTILE	MEAN (S.D.)	5TH PERCENTILE	95TH PERCENTILE
Wallerstein's prior						
Intercept	54.16 (34.61)	-2.76	111.09	31.19 (56.50)	-61.75	124.12
Left government	.29 (.06)	.19	.40	.29 (.07)	.18	.41
Logged labor-force size	-4.19 (2.12)	-7.67	70	-2.71 (3.57)	-8.58	3.16
Economic concentration	16.29 (13.02)	-5.12	37.70	23.72 (19.41)	-8.20	55.64
Stephens's prior						
Intercept	67.53 (19.01)	36.27	98.80	53.15 (38.93)	-10.88	117.18
Left government	.30 (.06)	.19	.40	.30 (.07)	.18	.41
Logged labor-force size	-4.88 (1.64)	-7.58	-2.17	-4.01 (2.64)	-8.34	.33
Economic concentration	10.87 (4.87)	2.86	18.87	15.88 (12.72)	-5.04	36.80

Note: In this analysis, the outlying case of Italy is omitted. The diffuse priors are the informative priors multiplied by 10 (N = 19). Least squares estimation with Italy excluded yields the following result (standard errors in parentheses): Union Density = 22.08 (63.11) + .30 (.07) × Left Government – 2.12 (4.00) × Size + 26.66 (21.41) × Concentration.

investigated with Bayesian regression diagnostics. Pettit and Smith (1986) discuss a statistic that indicates the influence of each sample observation (or set of observations) on the joint distribution of the posterior (see Appendix). This diagnostic indicates that Italy is highly influential under both informative priors. Both dimensions of the sensitivity analysissensitivity to the prior and sensitivity to the data—are shown in Figure 1. Here, influence statistics from the informative prior are plotted against influence statistics from the diffuse prior. If the sample data from a country have the same influence on the final results, regardless of the level of prior information, the country will fall on the solid 45-degree line. However, as the prior information is reduced, the sample data generally have an increasing impact on the posterior, so that the quantitative information from most countries becomes more influential as the prior information recedes. Italy grows disproportionately in its influence on the posterior. When prior information is eliminated, this country is shifted upward on the plot and indeed drives a large part of the story.

Table 5 summarizes the posterior distribution of the coefficients when Italy is omitted. These results show that Italy is influential for the posterior means of the size and concentration coefficients but has little impact on the posterior variances. Thus our posterior uncertainty changes little as a result of deleting Italy, but the range of values over which we are uncertain changes considerably. The new posterior distributions for the reduced data set indicate that Italy was driving up the posterior mean for the size effect but suppressing the posterior mean for concentration. Under Wallerstein's prior, the bulk of the posterior for the size coefficient remains negative, although the magnitude of the posterior mean declines by about a fifth. By contrast, the posterior mean for economic concentration increases by four times when Italy is omitted. Still, uncertainty about the effect remains substantial and the 90% confidence interval still includes zero.

Interestingly, Stephens's prior produces stronger evidence for the effect of size than Wallerstein's prior when Italy is omitted. This can be explained by noting that when there is strong dependency between two variables (size and concentration in this case), prior information for one can result in a sharper estimate for the effects of the other. In effect, an informative prior for the concentration coefficient frees up independent information for the size coefficient. (For a discussion and further examples, see Belsley 1991, 311, 318). This finding illustrates an implication of our discussion of collinearity: when variables are correlated, inference about one coefficient depends on prior information about others (Leamer 1978, 176). Like the size effect under Stephens's prior, the concentration effect remains fairly robust to the exclusion of Italy. Sensitivity analysis on the reduced data set also resembles the full sample analysis, indicating the crucial importance

of prior information for inference about the signs of the size and concentration effects. With diffuse priors in the reduced sample, all posterior size and concentration effects overlap zero.

CONCLUSION

We have tried to do two things. First, we have tried to identify two problems with conventional statistical inference in comparative research: weak and nonstochastic data. Second, we have outlined a Bayesian approach to statistical inference that offers unified and principled solutions to these disparate problems. Regardless of the merits of the Bayesian approach to statistical inference, the problems with conventional frequentist inference in comparative research that we identify will not go away. Conventional inference, like Bayesian inference, is based on a positive theory of how data are generated. This is not commonly recognized. In practice, data analysis also involves the use of extensive prior information. Again, this is not widely acknowledged. A chief concern of ours, then, is to problematize conventional statistical practice in an area for which conventional statistical inference was not particularly designed.

The Bayesian alternative that we propose is one way of addressing the problems of statistical inference in comparative research. It is based on a subjective probability concept that does not rely on data generated by a repeatable mechanism. The theory of data generation behind Bayesian inference is thus more general than the theory of data generation in conventional inference. Sources of uncertainty in the Bayesian approach are not restricted to—but could include-repetitious social processes. The other advantage of the Bayesian approach in the context of comparative research is the formal incorporation of prior information. When sample data are weak, prior information provides a useful supplement to the analysis. Still, the subjective choice of prior is an important weakness of Bayesian practice. The consequences of this weakness can be limited by surveying the sensitivity of conclusions to a broad range of prior beliefs and to subsets of the sample.

To summarize the data analysis, inferences about the negative effect of labor-force size and the positive effect of economic concentration on unionization are sensitive to the priors and the data. Analysis of the full sample showed that evidence for the size effect is consistent with a large set of informative and diffuse priors. Stephens's argument for a positive concentration coefficient, on the other hand, depends decisively on a prior belief in this effect. Italy drives confidence in the size effect and contributes uncertainty to the impact of economic concentration. When Italy is excluded from the analysis, all the results are weakened substantially, and no sign inference about either the concentration or size effect is sustained under the diffuse priors.

On balance, the sample data alone are not sufficiently informative to decide whether Wallerstein or Stephens is right. A resolution requires more information, which can be supplied through a prior distribution. The plausibility of the results is thus closely linked to the plausibility of the priors. Although this conclusion may appear unsatisfying, we would argue that it demands a lot from 20 data points to resolve a question of broad comparative scope concerning a complex process of institution building. More optimistically, if we allow information like Stephens's qualitative historical account of the impact of economic concentration on early Swedish unionization, a stronger set of results can be obtained. Thus, the data analysis underlines the importance of case studies and comparative histories to assist in the interpretation of quantitative evidence. While this is not news for comparative researchers, the Bayesian approach that we describe permits the transparent incorporation of this supplementary and influential information.

APPENDIX: BAYESIAN REGRESSION DIAGNOSTICS

Following the notation in the text, where s^2 is the sample estimate of the residual variance, Pettit and Smith's (1986) Bayesian influence statistic, I_i , expressing the influence of the *i*th observation on the joint posterior distribution of the regression coefficients, is

$$I_i = \frac{u_i}{2(1-u_i)} \left[u_i + \frac{e_i^2(2-u_i)}{s^2(1-u_i)} \right],$$

where $u = \text{diag}[s^{-2}(XV^*X')]$, and *e* is a vector of posterior residuals (the difference between the observed response and the fitted values from the observed data and posterior regression coefficients).

The *I*-statistic has parallels with conventional diagnostics. The matrix from which u is extracted can be understood as a "posterior hat-matrix" with the leverages of the sample observations on the posterior parameters along the main diagonal (Pettit 1985, 189; for discussion of leverage and the hat-matrix, see Cook and Weisberg 1982, 11–22). The scaled squared residuals in the expression for *I* can be thought of as squared "posterior studentized residuals" (Pettit and Smith 1986, 485).

S-Plus and GAUSS routines to calculate posterior regression parameters, confidence intervals, and Pettit and Smith's influence statistic are available on request. Press 1989 contains a useful software review.

Note

We thank David Gow for his comments on an earlier draft, and Richard Berk and Larry Bartels for their (seemingly unrelated) influences on our thinking about statistical inference. Chris Achen, Neal Beck, David Epstein, Don Green, Gary King, Sharyn O'Hallorhan, Marco Steenbergen and seminar participants at the University of California, Los Angeles, Department of Sociology provided helpful comments and criticisms. Earlier versions of this paper were presented at the annual meetings of the Midwest Political Science Association, Chicago, 1993 and the Tenth Annual Political Methodology Conference, Tallahassee, 1993.

References

- Alvarez, R. Michael, Geoffrey Garrett, and Peter Lange. 1991. "Government Partisanship, Labor Organization, and Macroeconomic Performance." American Political Science Review 85:539-56.
- Barnett, Vic. 1982. Comparative Statistical Inference. 2d ed. New York: Wiley.
- Barrilleaux, Charles J, and Mark E. Miller. 1988. "The Political Economy of State Medicaid Policy." American Political Science Review 82:1089-1107.
- Beck, Nathaniel, R. Michael Alvarez, Geoffrey Garrett, Jonathan Katz, and Peter Lange. 1993. "Government Partisanship, Labor Organization, and Macroeconomic Performance: A Corrigendum." American Political Science Review 87:945-948.
- Belsley, David A. 1991. Conditioning Diagnostics: Collinearity and Weak Data in Regression. New York: Wiley. Berk, Richard A., Bruce Western, and Robert Weiss. 1993.
- "Statistical Inference for Apparent Populations." University of California, Los Angeles. Typescript.
- Cochran, William G. 1953. Sampling Techniques. New York: Wiley
- Cook, R. Dennis, and Sanford Weisberg. 1982. Residuals and
- Influence in Regression. New York: Chapman & Hall. Efron, Bradley. 1986. "Why Isn't Everyone a Bayesian?" American Statistician 40:1–11.
- Erikson, Robert S., John P. McIver, and Gerald C. Wright Jr. 1987. "State Political Culture and Public Opinion." American Political Science Review 81:797-813.
- De Finetti, Bruno. 1974. Theory of Probability: A Critical Introductory Treatment. New York: Wiley. Fisher, Ronald A. 1935. The Design of Experiments. 1st ed.
- London: Oliver & Boyd. Fox, John. 1990. "Describing Univariate Distributions." In
- Modern Methods of Data Analysis, ed. John Fox and J. Scott Long. Newbury Park, CA: Sage.
- Freedman, D. A., and David Lane. 1983. "Significance Test-ing in a Nonstochastic Setting." In A Festschrift for Erich L. Lehman, ed. Peter J. Bickel, Kjell A. Doksum, and J. L.
- Hodges, Belmont, CA: Wadsworth.
 Golden, Miriam. 1993. "The Dynamics of Trade Unionism and National Economic Performance." American Political Science Review 87:439-54.
- Gorin, Zeev. 1980. "Income Inequality in the Marxist Theory of Development: A Cross-National Test." In Comparative Social Research, ed. Richard Tomasson. Greenwich, CT: 3A1.
- Granger, C. W. J., ed. 1990. Modelling Economic Series. Oxford: Clarendon.
- Hastie, Trevor J., and Robert J. Tibshirani. 1990. Generalized Additive Models. Chapman & Hall. London.
- Jackman, Robert W. 1985. "Cross-national Statistical Research and the Study of Comparative Politics." American Journal of Political Science 29:161-82.
- Kjellberg, Anders. 1983. Facklig Organisering i Tolv Länder. Lund: Arkiv.
- Lange, Peter, and Geoffrey Garrett. 1985. "The Politics of Growth: Strategic Interaction and Economic Performance in

the Advanced Industrial Democracies, 1974-1980." Journal of Politics 47:792-827.

- Lange, Peter, and Geoffrey Garrett. 1987. "The Politics of Growth Reconsidered." Journal of Politics 49:257-75.
- Leamer, Edward E. 1978. Specification Searches. New York: Wiley. Leamer, Edward E. 1983. "Let's Take the Con out of Econometrics." American Economic Review 23:31-43.
- Leamer, Edward E. 1991. "A Bayesian Perspective on Inference from Macro-economic Data." Scandinavian Journal of Economics 93:225-48.
- Lee, Peter M. 1989. Bayesian Statistics: An Introduction. New York: Oxford University Press.
- Pettit, L. I. 1985. "Diagnostics in Bayesian Model Choice." Statistician 35:183-90.
- Pettit, L. I., and A. F. M. Smith. 1986. "Outliers and Influential observations in Linear Models." In Bayesian Statistics 2, ed. J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith. Amsterdam: Elsevier Science.
- Pollard, William E. 1986. Bayesian Statistics for Evaluation Research: An Introduction. Beverly Hills: Sage.
- Powell, G. Bingham, Jr. 1982. Contemporary Democracies. Cambridge: Harvard.
- Press, S. James. 1989. Bayesian Statistics: Principles, Models and
- Applications. New York: Wiley. Radcliff, Benjamin. 1992. "The Welfare State, Turnout, and the Economy: A Comparative Analysis." American Political
- Science Review 86:444–54. Remmer, Karen L. 1991. "The Political Impact of Economic Crisis in Latin America in the 1980s." American Political Science Review 85:777-800.
- Robertson, John D. 1990. "Transaction-Cost Economics and Cross-national Patterns of Industrial Conflict: A Comparative Institutional Analysis." American Journal of Political Science 34:153-89.
- Stephens, John; Michael Wallerstein. 1991. "Industrial Concentration, Country Size, and Trade Union Membership. American Political Science Review 85:941–53. Swank, Duane H. 1988. "The Political Economy of Govern-
- ment Domestic Expenditure in the Affluent Democracies, 1960–1980." American Journal of Political Science 32:1120–50. Swank, Duane H. 1992. "Politics and the Structural Depen-
- dence of the State in Democratic Capitalist Nations." American Political Science Reviews. 86:38-54.
- Tanner, Martin A. 1993. Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions. 2d ed. New York: Springer-Verlag. Wallerstein, Michael. 1989. "Union Organization in Ad-
- vanced Industrial Democracies." American Political Science Review 83:481-501.
- Weede, E. 1986. "Sectoral Reallocation, Distributional Coalitions and the Welfare State as Determinants of Economic Growth Rates in Industrialized Democracies." European Journal of Political Research 14:501-19.
- Western, Bruce. N.d. "Unionization and Labor Market Institutions in Advanced Capitalist Countries, 1950-1985." American Journal of Sociology Forthcoming Wilensky, Harold L. 1981. "Leftism, Catholicism, Democratic
- Corporatism: The Role of Political Parties in Recent Welfare State Development." In The Development of Welfare States in Europe and America, ed. Peter Flora and Arnold J. Heidenheimer. New Brunswick: Transaction Books.
- Williams, John T. 1991. "The Political Manipulation of Macroeconomic Policy." American Political Science Review 84:765-95.

Bruce Western is Assistant Professor of Sociology, and a Faculty Associate at the Office of Population Research, Princeton University, Princeton, NJ 08544-1010.

Simon Jackman is Doctoral Candidate in Political Science, University of Rochester, Rochester NY 14627, currently visiting at the Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08544-1013.