Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach
Author(s): Bruce Western
Source: *American Journal of Political Science*, Vol. 42, No. 4 (Oct., 1998), pp. 1233-1259
Published by: Midwest Political Science Association
Stable URL: https://www.jstor.org/stable/2991856
Accessed: 14-07-2019 17:45 UTC

Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach

Bruce Western, *Princeton University*

Pooled cross-sectional time-series models in comparative politics typically constrain the effects of variables to be identical across countries. These models conflict with general principles of comparative analysis and theories of comparative political economy that the models are designed to test. In contrast, Bayesian hierarchical models allow time-series coefficients to vary across countries, and time-series effects can be related to cross-national variation in institutions. While allowing causal complexity into comparative analysis, the hierarchical model also provides: (1) more accurate forecasts than rival models; (2) more accurate estimates of time-series effects than unpooled analysis; and (3) a more realistic accounting of uncertainty than conventional pooled analysis. In addition, Bayesian theory for the hierarchical model helps specify the concept of "comparability" in comparative research. These ideas are illustrated in a reanalysis of a model of the political determinants of economic growth studied by Alvarez, Garrett, and Lange (1991).

A common design in comparative research combines time-series data from a number of countries for regression analysis. Applications often arise in comparative political economy where series of economic performance measures or policy outcomes are regressed on economic and political variables (e.g., Huber, Ragin, and Stephens 1995; Hicks 1994a; Swank 1992; Alvarez, Garrett, and Lange 1991; Korpi 1989). The popularity of the approach is also indicated by several recent papers discussing estimation of pooled cross-sectional time-series models (Stimson 1985; Beck and Katz 1995, 1996a; Hicks 1994b). For most of these analyses, the effects of causal variables are assumed to be identical across countries.

Bayesian hierarchical models provide an alternative approach in which cross-national variation in time-series coefficients is central to the analysis. In this approach, causation is heterogeneous. Consistent with the basic insights of comparative politics, hierarchical models monitor how political

Department of Sociology, 2-N-2 Green Hall, Princeton University, Princeton, NJ 08544-1010. E-mail: western@princeton.edu. This paper was motivated by persistent, skeptical, and generally difficult inquiries from Michael Shalev who also provided superb comments on an earlier draft. The paper also benefited from the excellent comments of Larry Bartels, Neal Beck, the editor, and anonymous reviewers. Research for this paper was supported by grant SBR95-11473 from the National Science Foundation. Michael Alvarez kindly supplied the data used for analysis, available from http://www.hss.caltech.edu/~rma/replication.html.

American Journal of Political Science, Vol. 42, No. 4, October 1998, Pp. 1233–1259

processes play out differently in different settings. In brief, Bayesian hierarchical models provide a way of pooling information from a set of countries to obtain optimal statistical estimates for any particular country. For the pooled cross-sectional time-series design, the models are hierarchical in the sense that country-level time-series coefficients are treated as if they were drawn from a population distribution shared by all the countries under analysis. In the Bayesian statistics that underpin the hierarchical model, the coefficients are described as "exchangeable." These country-level coefficients can then be written as a function of time-invariant institutional conditions. Unlike many quantitative analyses in comparative politics, hierarchical models thus present a rich picture of causal heterogeneity featuring institutions as an important source of contextual effects.

This specification has a strong substantive justification. Hierarchical models usefully capture Przeworski and Teune's (1970) classic account of contextual explanation for comparing political processes in different societies. By allowing greater causal complexity, the approach is similar in spirit to Ragin's (1987) emphasis on conjunctural causal explanation in comparative research. In comparative political economy, motivation for hierarchical models is particularly compelling. Here, institutions are viewed as enduring contexts which shape economic relationships. However, institutional explanation in this field is strongly at odds with typical analyses which impose a general causal story on all countries under study. With causal heterogeneity depending on cross-national variation in institutions, hierarchical models provide a closer fit between institutional theory and model specification.

While their substantive utility may be clear, hierarchical models also offer several statistical advantages. Although similar to conventional regressions with interaction terms, hierarchical models provide a more realistic assessment of uncertainty in comparative data. Interaction models assume that time-series effects depend deterministically on institutions. Hierarchical models relax this assumption by allowing time-series coefficients an extra stochastic component, an additional uncertainty which influences inferences. Inference for institutional effects reflects uncertainty about time-series coefficients which is ignored in interaction models. The time-series coefficients themselves are more accurate than estimates from unpooled data. The theory of hierarchical models also helps clarify general principles of comparative analysis by describing the "comparability" of countries. The platitudinous danger of comparing apples and oranges is given greater analytical content with the Bayesian idea of exchangeability.

I illustrate the hierarchical model in a reanalysis of the political determinants of economic growth in OECD countries studied by Alvarez, Garrett, and Lange (1991) and Beck et al. (1993). In contrast to the original analysis of Alvarez, Garrett, and Lange (1991), the hierarchical model admits substantial casual heterogeneity among the economic and political determinants of OECD economic growth. The model provides new evidence for how political institutions shape economic effects and generates more accurate forecasts of economic growth than rival models.

1. Comparative Politics and Causal Heterogeneity

The basic methodological divide in comparative research separates studies of generalizable, simple causes from analyses of conjunctural, complex causes (Ragin 1987). In one research tradition a general causal process is inferred with data from a very large number of countries. In the other, comparisons are small in scope, and local conditions significantly alter the causal story from place to place. Tilly (1984) and Lieberson (1992) critically review each approach. A middle course in comparative analysis emphasizes contextual explanation. Causal heterogeneity across countries is acknowledged, but a more general explanation of this heterogeneity is viewed as a central task for comparative research. A landmark statement of this position is detailed by Przeworski and Teune (1970), who believe that comparative research is fundamentally contextual, studying "the influence of larger systems upon the characteristics of units within them" (74). Ragin (1987) takes a similar position. He argues that conventional quantitative models fail to accommodate causal complexity, but comparative case studies do not provide a general vocabulary for describing cross-county variation. Ragin proposes a "qualitative comparative analysis" distinguishing cases according to different combinations of causal variables. In sum, both approaches offer contextual explanation as the characteristic method of comparative research. (see also Erbring and Young 1980; and Iversen 1991.)

From a substantive point of view, the role of institutions in comparative political economy provides a natural motivation for contextual theories. Schumpeter (1954, 34) provides an early statement of this idea, claiming that " 'economic laws' are much less stable than are the 'laws' of any physical science. . . . [T]hey work out differently in different institutional conditions, and . . . the neglect of this fact has been responsible for many an aberration." This idea has been adopted in modern comparative studies of economic performance. For instance, Bruno and Sachs (1985, 274) argue that general economic theories wrongly "consider one and the same basic model as applicable to all economies." They argue, instead, that the logic of market processes within national economies is modified by labor market institutions. In relation to unemployment, Solow (1986, S23) makes a similar observation: "more often than not we fail to take institutional differences seriously. One model is supposed to apply everywhere and always. Each country is a just a point on cross-section regression or one among several essentially identical regressions. . . ." For these institutionally-oriented economists, economic relationships depend on institutional contexts.

Comparative political scientists studying economic processes have also examined the contextual effects of institutions. A leading example is the work of Lange and Garrett (1985; Alvarez, Garrett, and Lange 1991). Endorsing the institutional argument of Bruno and Sachs, Lange and Garrett argue that the impact of political parties on macroeconomic performance is conditional on the development of corporatist institutions for centralized collective bargaining. In their statistical analysis, these conditional effects were modelled with interaction terms. The importance of contextual variation in political processes is discussed more generally in the historical institutionalism of Thelen and Steinmo (1992). Their research program has targeted a range of outcomes for contextual explanation, including trade union growth, taxation policy, and industrial relations. For much of this research, institutions are not simply another variable in a list of explanatory factors; instead their influence is pervasive, "mediating relations of cooperation and conflict, . . . structur[ing] political situations and leav[ing] their own imprint on political outcomes" (Thelen and Steinmo 1992, 9). In the following section I describe how the pervasive and contextual influence of institutions can be captured with hierarchical models.

2. Introducing the Hierarchical Model

Although quantitative comparative research often emphasizes the impact of institutions and uses the pooled cross-sectional time-series research design, the full implications of institutional explanation are rarely investigated. The dominant approach fits a linear regression of the following form:

$$y_{it} = \beta_0 + \beta_1 x_{it} + e_{it}, \tag{1}$$

where the dependent variable, y in country i ($1 = 1, \ldots, I$) at time t ($t = 1, \ldots, T_i$) is a linear function of predictors, in this case x. The coefficients, β, are the same for all countries, indicating that causal processes relating x to y show no cross-national variation. Variation across countries is sometimes acknowledged in structure placed on the errors, e, that allows for unequal residual variance and cross-country correlations. The serial character of the data is often addressed in an adjustment for residual first-order autoregression (Hicks (1994b) reviews these specifications). A covariance matrix reflecting this error structure can be estimated and used in generalized least squares (GLS) to estimate coefficients. In this case the cross-sectional and serial character of the data offer no special substantive insight, and variation arising out of the data structure is relegated to the error term. With these models, the research design only contributes additional data points, providing greater statistical precision for estimates of the coefficients, β_0 and β_1.

To allow for the possibility of contextual variation, the model for a single country can be rewritten:

$$y_{it} = \beta_{0i} + \beta_{1i} y_{it-1} + \beta_{2i} x_{it} + e_{it}. \tag{2}$$

Unlike Equation 1, the dynamics of the response variable are introduced into the structural part of the equation through the lagged dependent variable (see Beck and Katz 1996a). The coefficients also carry the i subscript indicating that the regression relationships vary across countries. Variation in the time-series coefficients is then written as a function of time-invariant institutional conditions,

$$\begin{aligned} \beta_{0i} &= \alpha_{00} + \alpha_{01} z_i + \eta_{0i}, \\ \beta_{1i} &= \alpha_{10} + \alpha_{11} z_i + \eta_{1i}, \\ \beta_{2i} &= \alpha_{20} + \alpha_{21} z_i + \eta_{2i}, \end{aligned} \tag{3}$$

The subscripts on the α coefficients indicate that the institutional effects are constant across countries. Here and throughout I refer to time-invariant variables as "institutional." From a methodological viewpoint, of course, any time-invariant variable—institutional or otherwise—could be used in the analysis. In some real applications, institutional variables may move slowly over time. This type of variation is difficult to accommodate in the current set-up, although analysis of models with institutional change can usefully be approached from a Bayesian perspective (Western 1997a).

The hierarchical model can be written as a single equation with interaction terms by substituting Equation 3 into Equation 2:

$$\begin{aligned} y_{it} &= \left(\alpha_{00} + \alpha_{01} z_i + \eta_{0i}\right) + \left(\alpha_{10} + \alpha_{11} z_i + \eta_{1i}\right) y_{it-1} \\ &\quad + \left(\alpha_{20} + \alpha_{21} z_i + \eta_{2i}\right) x_{it} + e_{it} \\ &= \alpha_{00} + \alpha_{01} z_i + \alpha_{10} y_{it-1} + \alpha_{20} x_{it} \\ &\quad + \alpha_{11} z_i y_{it-1} + \alpha_{21} z_i x_{it} + \left(\eta_{0i} + \eta_{1i} y_{it-1} + \eta_{21} x_{it} + e_{it}\right) \end{aligned}$$

This is identical to the usual single-equation regression with interaction terms except the error term has a more complicated structure including two sources of uncertainty, e and η. From a non-Bayesian perspective, this specification is sometimes called a "random effects" model because the time series coefficients have a stochastic component, η. We can think of these stochastic components as drawn from a single population distribution that is shared by all the countries under study. The institutional effects—the "fixed effects"—describe how institutions shape the time-series relationship unfolding within each country.

The institutions also exert a direct influence on the general level of dependent variable by driving variation in the time-series intercepts, β_{0i}. An important transformation centers all the time-series predictors, including the lagged dependent variable. For instance, x_{it} is replaced by $x_{it} - \bar{x}_i$, where $\bar{x}_i$ is the country mean of x_{it}. By centering all the predictors, the intercepts for each country from unpooled analysis equal the country means of the dependent variable. The transformation thus incorporates a substantively interesting cross-sectional analysis into the pooled cross-sectional design. The institutional effects for the country-level intercepts describe the impact of institutions on the average level of the dependent variable.

The interaction model is obtained by setting the random effects to zero, $\eta = 0$. Equation 3 shows that this specification is equivalent to assuming that the institutional conditions can predict variation in the time-series coefficients without error. Because such deterministic relationships rarely arise in applied contexts, hierarchical models provide a more realistic account of all the sources of uncertainty associated with the comparative research design than the model with interaction terms.

The stochastic terms, η, affect estimation of the institutional effects, α, and the times-series effects, β. To see this, I generalize the notation by collecting the K time-series predictors for country i in the $(T_i \times K)$ matrix, X_i, and the dependent variable in the vector, y_i.

Following the discussion of Bryk and Raudenbush (1992, 37–38), the time-series equation is written:

$$y_i = X_i\beta_i + e_i, \tag{4}$$

Premultiplying both sides of Equation 4 by $\left(X_i'X_i\right)^{-1}X_i'$ gives

$$\hat{\beta}_i = \beta_i + \nu_i, V\left(\nu_i\right) = V_i, \tag{5}$$

where the country-level ordinary least squares estimator, $\hat{\beta}_i = \left(X_i'X_i\right)^{-1}X'y_i, \nu_i = \left(X_i'X_i\right)^{-1}X_i'e_i, V_i = \sigma_i^2\left(X_i'X_i\right)^{-1}$ is the variance of the least squares estimator, and σ_i^2 is the residual variance of y_i. Equation 4, describing the effects of the time-series variables within countries, is sometimes called the micro model.

If the J institutional variables for country i are collected in $(1 \times J)$ row vector, z_i', the block diagonal $K \times (J \times K)$ matrix $Z_i = \text{diag}(z_i, \ldots, z_i)$. The time series coefficients then depend linearly on Z_i,

$$\beta_i = Z_i\alpha + \eta_i, V(\eta_i) = U, \tag{6}$$

with zero-expectation random effects, η_i. Equation 6, describing the effects of institutional variables across countries, is called the macro-model. The $(K \times K)$ covariance matrix, U, is constant across countries and indicates the dispersion of the time-series coefficients around the expected value, $Z_i\alpha$. Substituting Equation 6 into Equation 5 gives

$$\hat{\beta}_i = Z_i\alpha + \eta_i + \nu_i,$$

where the variance of $\hat{\beta}_i$ conditional on the institutional variables, Z_i, is

$$V(\hat{\beta}_i) = V(\eta_i + \nu_i) = U + V_i = \Delta_i.$$

The variance of the estimated time-series coefficients thus have two components that will, in general, vary across countries. There is variance attributable to uncertainty at the micro-level within countries, V_i; there is also variance due to uncertainty at the macro-level across countries, U. Both levels of uncertainty contribute to estimates of the institutional effects with the GLS estimator:

$$\hat{\alpha} = \left(\sum_i Z_i'\Delta_i^{-1}Z_i'\right)^{-1} \sum_i Z_i\Delta_i^{-1}\hat{\beta}_i, \tag{7}$$

with variance

$$V(\hat{\alpha}) = \left(\sum_i Z_i'\Delta_i^{-1}Z_i'\right)^{-1}.$$

The variance of the institutional effects will depend partly on V_i that measures how precisely the time-series data estimate β_i within each country. In practice the addition of random effects, η, tends to inflate uncertainty about the institutional effects compared to naive estimates that simply regress $\hat{\beta}_i$ on the institutional data. (The comparative study of Griffin, O'Connell, and McCammon 1989 provides an example of regressions on country-level regression coefficients.) Overconfidence in institutional effects in the naive regression results from the false assumption that the time-series coefficients have been estimated without error.

The time-series coefficients, β, themselves will also be of substantive interest. Two alternative estimates are suggested by the previous discussion:

$$\hat{\beta}_i = (X_i'X_i)^{-1}X_i'y_i \text{ and } \overline{\beta}_i = Z_i\alpha.$$

We can think of $\hat{\beta}_i$ as the micro-level estimate using the time-series data and $\overline{\beta}_i$ as the macro-level estimate based on the institutional data. A compromise estimator is the matrix-weighted sum of the micro- and macro-level estimates:

$$\beta_i^* = W_i\hat{\beta}_i + (I - W_i)\overline{\beta}_i,$$

where $W = V_i(V_i + U)^{-1}$, and I is the identity matrix conformable with W. Under distributional assumptions described next, β_i^* is the Bayesian posterior expectation of β_i. If the data from each country yield sharp least squares estimates, V_i is small compared to U and $\hat{\beta}_i$ will dominate the final result. On the other hand, if $\hat{\beta}_i$ has large variance but is accurately predicted by the institutional data, β_i^* will be close to $\overline{\beta}_i$.

The Bayes estimate, β_i^*, has several useful statistical properties. Under quite general conditions, the Bayes estimate will have smaller mean squared error than the least squares estimate, $\hat{\beta}_i$ (Lindley and Smith 1972, 3). This increased accuracy can be intuitively understood to result from the additional data used to estimate U and $Z_i\alpha$ upon which β_i^* is based. The Bayes estimates for each context are sometimes described as "borrowing strength" from the whole sample (Lindley and Smith 1972). In the extreme case of borrowing strength, the Bayes estimate for a particular country is estimated even if the country-level time-series predictors, X_i, exhibit perfect collinearity. Information from other countries will help provide an estimate for a coefficient in a particular country where, say, a given independent variable shows no variation. In this case the Bayes estimate approximately equals its expected value, $\beta_i^* \approx Z_i\alpha$. This is important in comparative applications where country-level intercepts are usually unidentified in the presence of time-invariant institutional variables. With the hierarchical model, country-level intercepts can be introduced as random effects even with variables that show no variation over time.

The Bayes estimate is also called a shrinkage estimator because it shrinks the coefficients back to their expected value, $Z_i\alpha$. The shrinkage effects causes the time-series coefficients to be more tightly clustered than those from regressions run separately for each country. As we'll see below, this shrinkage behavior also yields superior out-of-sample predictions compared to results from regressions run separately for each country.

Random effects models have a long history in statistics and econometrics. Beck and Katz (1996b) have recently provided a good discussion of random coefficient models for comparative politics. Their model is a special case of the current specification, omitting institutional effects and the issue of contextual explanation. Good accessible surveys of hierarchical models include Bryk and Raudensbush (1990), Mason, Wong, and Entwistle (1983), Robinson (1991), and de Leeuw and Kreft (1986). Likelihood and frequentist inference dominates the discussion in most of these sources, however. In the following section, I review some Bayesian aspects of the hierarchical model.

3. THE BAYESIAN HIERARCHICAL MODEL

3.1 Model Specification

Suppose we have a dependent variable y_{it} from country i and time t that depends linearly on a vector of predictors x_{it} that may include lagged values of the dependent variable. The time-series coefficients, β_i, vary across countries and linearly depend on a time-invariant $(J \times 1)$ vector of institutional conditions, z_i. Commonly y_{it} and β_i are assumed to be conditionally normal:

$$\begin{aligned} y_{it} &\sim N(\hat{y}_{it}, \sigma_i^2), i = 1, \ldots, I \text{ and } t = 1, \ldots, T_i, \\ \hat{y}_{it} &= x_{it}'\beta_i. \end{aligned}$$

In common with the usual pooled-cross-sectional time series models, this model allows for cross-country heteroskedasticity.

If coefficients from all I countries for time series predictor k ($k = 1, \ldots, K$) are collected in the $(I \times 1)$ vector, β_k, the stochastic character of the time-series coefficients is written

$$\begin{aligned} \beta_k &= Z\alpha_k + \eta_k, \\ \eta_k &\sim N(0, \tau_k^2), \end{aligned} \tag{8}$$

where the institutional variables for all countries are collected in the $(I \times J)$ matrix, Z. This distribution for the η coefficients is called a population distribution (Gelman et al. 1995, 377), and the parameters α_k and τ_k^2 are called hyperparameters. This notation demonstrates the hierarchical nature of the model. The dependent variables is drawn from a sampling distribution characterized by parameters. The parameters are drawn from a population distribution characterized by hyperparameters.

To complete the model, conjugate priors are given for the hyperparameters:

$$\begin{aligned} \sigma_j^{-2} &\sim \text{Gamma}\left(m_\sigma, n_\sigma\right) \\ \tau_k^{-2} &\sim \text{Gamma}\left(m_\tau, n_\tau\right) \\ \alpha_k &\sim N(p, q) \end{aligned} \tag{9}$$

The priors allow for a proper Bayesian analysis, but in practice they are usually taken to be diffuse and do not contribute importantly to the inferences of interest. In a comparative setting, knowledge about specific countries could be introduced through the prior distributions. Application of informative priors in Bayesian comparative research has been suggested by Western and Jackman (1994) and Bartels (1996).

So far the hierarchical model has: (1) provided a useful way of describing contextual effects in pooled cross-sectional time-series models, (2) suggested an estimator of these effects that takes account of uncertainty in the time-series coefficients, and (3) indicated a strategy for obtaining time-series coefficients that are more accurate than conventional lest squares estimates from unpooled data. While these issues are of immediate practical importance, the hierarchical model also sheds light on some general principles of comparative analysis.

In Bayesian statistics the stochastic character of the time series coefficients, β_i, is described as exchangeable. Formally, the joint probability distribution $p(\beta_1, \ldots, \beta_I)$ is invariant under permutations of the indices $(1, \ldots, I)$ (Gelman et al., 1995, 123–24). In a comparative setting this means that the country names—the indices of comparative research—carry no information about the time series coefficients in addition to the predictors. Consider Sweden (S) and the United States (A). If coefficients for these two countries are exchangeable, there is no reason to think that deviations of the Swedish coefficient from its expected value under the macro model, η_S, will be larger or smaller than the U.S. deviation, η_A. Exchangeability describes the comparability of countries. All relevant information about the countries is carried in the sample data, and residual variability in the statistics of interest is unrelated to the country names. While exchangeability describes prior indifference about the relative size of statistical parameters, it can also be interpreted as a claim about the quality of the model specification. If the β coefficients are exchangeable, the model explaining variation in β has taken account of all systematic structure in the coefficient distribution.

Przeworski and Teune (1970, 11) propose a similar idea of comparability as measurement. In their formulation, two countries are comparable if they can be described "in terms of a standard language or, simply [by] . . . *measuring* them (Przeworski and Teuene 1970, 11; original emphasis). The measurement concept of comparability answered the basic objection that social systems cannot be compared because they are unique. This objection is summarized in warnings against comparing "apples and oranges." By defining comparability in terms of measurement, "the response to the classic objection of comparing 'apples and oranges' is simple: they are 'fruits' " (Przeworski and Teune 1970, 10). However, the measurement concept disposed of one conceptual chestnut but left another.

If comparability is established by measurement, how do we respond to the familiar reproach that a measure—say the rate of economic growth—"means" something different in one country compared to another. This vaguely defined difference in meaning can be made more precise with the idea of exchangeability. If the coefficients describing economic growth in country A and country B are nonexchangeable, growth in the two countries is described by different models. Measurement in a common metric doesn't guarantee the comparability of two countries. Instead this depends on their description as particular realizations of a single causal model. In the Bayesian approach then, the comparability of two countries depends on their exchangeability: the researcher's belief that information from both countries is produced by the same causal process. (The idea of exchangeability embodying "comparability" or "similarity" judgments is discussed by Draper et al., 1993.)

3.2 Estimation

The hierarchical model can be estimated with the Gibbs sampler. The Gibbs sampler is a method for simulating draws from the posterior distribution of the time-series coefficients and the institutional effects. By generating many such draws, summaries of the posterior distribution can be calculated. In the application below, the Gibbs sampler is run for 10,000 iterations, yielding 10,000 sets of time-series coefficients and institutional effects. An estimate for an institutional effect, for example, is calculated by taking the mean or median from this simulated sample of size 10,000. Confidence intervals or standard errors of the estimates are similarly calculated from percentiles or standard deviations of the sampled effects.

Although a detailed description of the Gibbs sampler is beyond the scope of this paper, its basic features in this context can be simply described. Assume for now that all the variance parameters are known with certainty. With the distributional assumptions described above, the time series coefficients are normally distributed, given the sample data and the institutional effects. The institutional effects are also normally distributed, conditional on the data and the time-series coefficients. We are interested in simulating from the unconditional distribution of time series and institutional effects. Such simulation proceeds as follows:

1. Randomly generate a set of time-series coefficients, β_1, from a normal distribution, conditioning on the sample data, and some simulated values for the institutional effects, α.
2. Randomly generate a set of institutional effects, α, from a normal distribution, conditioning on the sample data and the time-series coefficients, β, generated in step 1.

Repeating steps 1 and 2 numerous times ultimately yields draws from the unconditional posterior distribution of α and β. In practice additional sampling steps must also be introduced to obtain variance estimates, but the underlying idea is the same. Random draws of one set of parameters are generated from known conditional distributions. These draws are used to update the conditional distributions of other parameters in the model.

An accessible account of the statistical theory for the Gibbs sampler is provided by Casella and George (1992). Tanner (1993, esp. 108–110) provides a more general discussion and details a variety of applications. The Gibbs sampler and related methods for hierarchical linear models are described by Gelman et al. (1995, 378). These sources illustrate the flexibility of the Gibbs sampler. For instance, the algorithm is readily extended to robust analysis with outliers (e.g., Western 1997b, chap. 7). In this case a heavy-tailed distribution like the *t* distribution replaces the normal sampling distribution or the normal population distribution.

Software for hierarchical models is now widely available. An extensive review of five packages for hierarchical modelling is reported by Kreft, de Leeuw, and van der Leeden (1994). In addition to specialized software, routines can also be found in the general statistical software SAS and S-Plus. Posterior simulation using the Gibbs Sampler is implemented in the BUGS program (Spiegelhalter et al., 1996). BUGS is particularly useful because it can be used to fit a wide variety of hierarchical models, and the software is free.

4. Institutions and OECD Economic Growth

To illustrate the hierarchical model I reanalyze data on OECD economic growth studied by Alvarez. Garrett, and Lange (1991). Like the discussion here, Alvarez, Garrett and Lange (1991) viewed time-invariant institutions as an important source of contextual effects. Their study linked macroeconomic outcomes to the strength of leftist parties in government. For their institutional theory, the economic effect of leftist government was conditional on the strength of labor unions. Centralized and encompassing unions helped leftist governments promote economic growth. Weak unions diminished the positive effects of leftist governments.

Table 1. Means of Time Series Variables, and Scores for Labor Organization Used for Analysis of Economic Growth, Sixteen OECD Countries, 1971–1984

	GDP Growth	OECD Demand	OECD Export	OECD Import	Leftist Government	Labor Organization
Australia	3.06	90.18	261.24	257.47	32.71	1.77
Austria	3.03	203.78	584.06	578.71	97.71	3.19
Belgium	2.34	344.48	990.89	973.70	24.00	2.77
Canada	3.69	136.63	396.85	388.13	0.00	0.98
Denmark	2.30	180.08	518.14	509.88	68.07	2.81
Finland	3.32	164.86	480.53	473.18	45.21	2.80
France	2.91	119.69	348.46	342.17	23.64	0.81
Germany	2.32	139.61	403.47	400.14	62.07	1.73
Ireland	3.25	293.50	845.13	836.40	21.07	1.80
Italy	2.39	136.38	397.16	392.33	12.64	1.58
Japan	4.58	68.20	208.01	202.65	0.00	0.41
Netherlands	2.18	298.41	825.94	819.62	16.71	1.89
Norway	4.17	250.10	712.04	701.65	69.00	3.46
Sweden	1.91	165.60	485.65	473.88	57.14	3.62
United Kingdom	1.71	156.82	448.50	439.88	36.86	1.93
United States	2.86	46.60	140.18	137.46	0.00	0.86

Note: Alvarez, Garrett, and Lange (1991) describe the data and their original sources.

4.1 Data and Model

Descriptive statistics for the data are reported in Table 1. Economic growth is measured by the percentage growth in real Gross Domestic Product (GDP), and data were collected for sixteen OECD countries from 1971 to 1984. Independent variables included a longitudinal measure of leftist government and a time-invariant, institutional variable indexing the strength of labor organization. Leftist government was measured by the percentage of cabinet seats held by leftist parties. A scale for labor organization combined information about union density and union centralization. Four other variables controlled for surrounding economic conditions. The influence of the domestic economy was captured by a lagged growth variable. International economic conditions were controlled by variables measuring vulnerability to demand in the OECD area and price movements of OECD imports and exports.

To model the conditional effect of labor parties, Alvarez, Garrett, and Lange (1991) introduced an interaction between leftist government and labor organization. For country i ($i = 1, \ldots, 16$) at time t ($t = 1971, \ldots, 1984$) the basic model is written:

$$y_{it} = \beta_0 + \beta_1 y_{it-1} + \beta_2 D_{it} + \beta_3 I_{it} + \beta_4 E_{it} + \beta_5 L_i + \beta_6 G_{it} + \beta_7 L_i \times G_{it} + e_{it}, \tag{10}$$

where y_{it} is economic growth, D_{it} is OECD demand, I_{it} is price movements of OECD imports, E_{it} is price movements of OECD exports, L_i is labor organization, G_{it} is leftist government, and e is an error term. With this model, only the effect of leftist government differs across countries.

Is it realistic to assume with Alvarez, Garrett, and Lange (1991) that economic effects are identical across countries? The institutional perspective that motivated the interaction model for leftist government can also apply to the economic variables. The authors recognize this, citing macroeconomic research which finds that that economic effects vary cross-nationally, depending on the labor market institutions (Alvarez, Garrett, and Lange 1991, 541). Among these authorities, for example, Bruno and Sachs (1985, 224) argue that highly encompassing unions tailor wage agreements to maintain the competitiveness of export industries. As a result, the damaging effect of international economic forces is reduced where labor organization is high. Research in comparative politics similarly suggests that centralized labor organization can dampen the negative effects of international market conditions on macroeconomic performance (e.g., Katzenstein 1985; Scharpf 1991). Thus economic effects, like the effects of partisan politics, may depend on the level of labor organization.

This discussion suggests two models that allow economic and political effects to vary across countries. Building on Alvarez, Garrett, and Lange (1991), interactions between the economic variables and labor organization could be added to Equation 10:

$$\begin{aligned} y_{it} = {} & \beta_0 + \beta_1 y_{it-1} + \beta_2 D_{it} + \beta_3 I_{it} + \beta_4 E_{it} + \beta_5 L_i + \beta_6 G_{it} \\ & + \beta_7 L_i \times y_{it-1} + \beta_8 L_i \times D_{it} + \beta_9 L_i \times I_{it} \\ & + \beta_{10} L_i \times E_{it} + \beta_{11} L_i \times G_{it} + e_{it} \end{aligned} \tag{11}$$

A complicated structure could be assumed for the errors, but Beck et al. (1993) find little difference between the constant variance model estimated with OLS and a more general heteroskedasticity model. I thus assume that

$$y_{it} \sim N(\hat{y}_{it}, \sigma^2)$$

and report the OLS results.

I compare the interaction model of Equation 11 to a similar Bayesian hierarchical model that also allows the time-series effects to depend on labor organization. For this model, time-series relationships within countries are written in the micro-level equation:

$$y_{it} = \beta_{0i} + \beta_{1i}y_{it-1} + \beta_{2i}D_{it} + \beta_{3i}I_{it} + \beta_{4i}E_{it} + \beta_{5i}G_{it} + e_{it}. \qquad (12)$$

The time-series coefficients carry the subscript i indicating that the longitudinal effects vary across countries. Information about labor organization enters through the macro-level equation which explains cross-national variation in the country-level time-series coefficients:

$$\begin{aligned} \beta_{0i} &= \alpha_{00} + \alpha_{01}L_i + \eta_{0i} \\ &\vdots \\ \beta_{5i} &= \alpha_{50} + \alpha_{51}L_i + \eta_{5i} \end{aligned} \qquad (13)$$

The hierarchical model is based on a normal distribution for the dependent variable:

$$y_{it} \sim N\left(\hat{y}_{it}, \sigma_i^2\right)$$

where the expectation, $E\left(y_{it}\right) = \hat{y}_{it}$ is now taken from Equation 12. If the sixteen random effects for k ($k = 0, \ldots, 5$) are collected in the vector η_k, then

$$\eta_k \sim N\left(0, \tau_k^2 I\right)$$

where I is the (16×16) identity matrix. The Bayesian model is completed with proper priors for the hyperparameters:

$$\begin{aligned} \sigma_i^{-2} &\sim \text{Gamma}(.001, .001) \\ \tau_k^{-2} &\sim \text{Gamma}(.001, .001) \\ \alpha_{jk} &\sim N(0, 10000), \end{aligned}$$

where j indexes the institutional predictors. These priors are quite flat and do not greatly effect the analysis.

The hierarchical model and the interaction model are identical except for the random effects, η. While the interaction model contains just one source of uncertainty, the error e, the hierarchical model has two, e and η. The random effects indicate that labor organization does not perfectly predict the time-series coefficients. There is additional variation beyond that explained by the institutional variable.

The macro-level Equation 13 could be revised by eliminating some of the labor organization effects. If labor organization only predicted the time-series intercepts, β_{0i}, and the leftist government effects, β_{5i}, a hierarchical version of the original Equation 10 would result. Under this model, the economic effects, $(\beta_1, ..., \beta_4)$, could still vary across countries, but independently of the level of labor organization. Still, the complete interaction model and its hierarchical counterpart have an interesting substantive motivation, and I focus on these models in the following analysis.

For both models, the time-series predictors are centered within countries. With this transformation, the country-level means of time-series predictors are zero. The institutional effect on the time series intercepts then describes the impact of labor organization on average growth.

Before examining the results, I consider two kinds of variation generated by the pooled cross-sectional time-series design: heteroskedasticity over time, and across countries. To allow shifts in the error variance over time, Beck et al. (1993) included dummy variables indicating each year of the time series. To control cross-sectional heteroskedasticity, the authors discussed a model with country-level intercepts. Such intercepts could not be estimated, however, because the time-invariant institutional variable was linearly related to the country dummies. The authors settled on a general heteroskedasticity estimator that adjusted the standard errors of the conventional OLS results.

The hierarchical model adapts easily to these kinds of variation. First, because Bayesian pooling borrows strength from the whole sample, linear dependencies in the data within countries present no obstacles to estimating country-level intercepts. Second, the likelihood could also allow the error variance to vary across countries. Rather than assume the variances are known as in estimated GLS, the unconditional posterior distributions used for Bayesian analysis integrates over uncertainty about those error variances. By accounting for uncertainty about variance parameters in inferences about coefficients, Bayesian analysis is thus more conservative than the original GLS estimates of Alvarez, Garrett, and Lange (1991). Third, for simplicity, I ignore the issue of heteroskedasticity over time. In the following analysis the focus is on forecasting growth rates beyond the observed time period. This exercise presents difficulties for models of longitudinal heteroskedasticity that are beyond the scope of this paper. If there is strong heteroskedasticity over time, there will be some efficiency loss in the current analysis, and inferences will be biased in a conservative direction.

4.2 Results

Regression results for the hierarchical and interaction models are reported in Table 2. The hierarchical model, fitting country-level slopes and intercepts, has many more parameters than the interaction model. As a re-

Table 2. Interaction Effects and Hierarchical Institutional Effects in a Model of Economic Growth, Sixteen OECD Countries, 1971–1984

	Hierarchical Model		Interaction Model	
	Labor Intercept	Labor Organization	Intercept	Organization
Intercept	3.264	–.218	3.317	–.218
	[2.730, 3.821]	[–.474, .030]	[2.919, 3.715]	[–.396, –.040]
Lagged Growth	.090	–.063	.158	–.054
	[–.079, .268]	[–.148, .021]	[–.018, .334]	[–.135, .028]
OECD Demand	.034	–.009	.019	–.005
	[.017, .050]	[–.016, –.002]	[.013, .025]	[–.007, –.002]
OECD Exports	–.001	.001	.000	.001
	[–.011, .009]	[–.003, .006]	[–.002, .003]	[.000, .002]
OECD Imports	–.008	.001	–.004	.000
	[–.023, .007]	[–.006, .007]	[–.011, .003]	[–.003, .003]
Leftist Government	–.022	.018	–.018	.011
	[–.065, .020]	[.000, .036]	[–.034, –.001]	[.004, .018]
R^2	.65		.35	

Note: Hierarchical model effects are coefficients from Equation 13. Interaction model effects are the main effects and interaction effects of Equation 11.

sult, the hierarchical model explains nearly twice the variance of the interaction model and a majority of the variance in the dependent variable. This sharp difference in goodness-of-fit suggests there is substantial parameter heterogeneity neglected by the simpler interaction model.

The remainder of Table 2 reports coefficients from the two models. In the Bayesian approach of this paper, I summarize the posterior distribution of the coefficients with an 80 percent confidence interval instead of providing a *p*- value or *t* statistic for a frequentist hypothesis test. (Analogous statistics could be calculated, however.) Focusing on the hierarchical model, the labor organization coefficient for the country-level intercepts shows the impact of unions on average growth rates. The negative coefficient indicates that economic growth is generally slower in countries with strong unions. For example, Sweden scores 3.62 on the labor organization scale, while Japan scores .41. This difference in labor organization is estimated to account for about $.70 = -.218 \times (.41\text{–}3.62)$ of a percentage point difference in average growth rates. Still, the confidence interval for the coefficient just overlaps zero, so we cannot be confident of a negative labor organization effect.

The original analysis of Alvarez, Garrett, and Lange (1991) focused on the conditional effects of leftist government. The hierarchical model replicates the finding of the earlier study. The positive institutional coefficient indicates that high labor organization allows leftist governments to boost economic growth. When labor organization is low, the effect of leftist government is close to zero. A one point increase in labor organization—roughly the difference between France and Germany—is estimated to raise the impact of leftist government by about 1.5 percentage points on the growth rate. The confidence interval excludes zero, providing strong support for the effect. Because the hierarchical model admits uncertainty at the micro- and macro-levels, inferences for the institutional effects tend to be conservative compared to interaction models. Despite this added uncertainty in the Bayesian model, there is still strong evidence that leftist governments and strong unions together promote economic growth.

While the hierarchical analysis of partisan effects is in line with earlier results, the model also provides novel evidence of how economic effects are patterned by institutions. The data analysis strongly supports the impact of labor organization on the effects of OECD demand. The negative labor organization coefficient indicates that the link between international demand and growth is weaker in countries with strong unions. Consistent with theory, this suggests that a high labor organization insulates economies from international market forces. When OECD demand declined dramatically in 1975 and again in 1982, countries with encompassing unions more successfully maintained economic growth relative to their counterparts with weak labor movements.

The interaction model provides similar results to the hierarchical model. The interaction model returns a significant positive coefficient for the influence of labor organization on leftist government. The model also yields a significant interaction between labor organization and OECD demand. The interaction effects are usually smaller than the hierarchical institutional effects, but confidence intervals for the interaction effects are about two-thirds the length of those from the rival model. Weaker inferences for the Bayesian institutional effects results from a more realistic assessment of uncertainty in which labor organization does not exhaustively explain all variation in the time series effects.

The institutional effects suggest substantial variation in the country-level coefficients for OECD demand and leftist government. More detail is given by the Bayes shrinkage estimates of the time-series coefficients (Table 3). In contrast to the assumption of invariant economic effects in the analyses of Alvarez, Garrett and Lange (1991) and Beck et al. (1993), the economic coefficients show striking variation across countries. Outside of Europe, the shrinkage coefficients indicate that growth rates are highly sensitive to OECD demand. This suggests that the non-European economies were more affected by the falls in international demand through the 1970s

Table 3. Bayes Shrinkage Estimates of Time-Series Effects in a Hierarchical Model of GDP Growth, Sixteen OECD Countries, 1971–1984

	Average Growth	Lagged Growth	OECD Demand	OECD Exports	OECD Imports	Leftist Government
Australia	2.98*	–.01	.21*	.01	–.03	.09
Austria	2.84*	–.08	.06*	.04*	–.10*	.63*
Belgium	2.49*	–.12	.02	.04*	–.09*	.25
Canada	3.43*	.03	.20*	.02	–.04	–.04
Denmark	2.42*	–.13	.06	.03	–.16*	.18*
Finland	2.94*	–.06	.09	.04*	–.03	.56*
France	2.95*	.02	.06	.04*	–.16*	–.27*
Germany	2.43*	–.03	.19*	.02	–.05	.29*
Ireland	3.02*	.00	.05	.01	–.01	.03
Italy	2.61*	–.01	.26*	.07*	–.04	–.02
Japan	3.94*	.07	.27*	–.06	–.02	–.14
Netherlands	2.40*	.01	.07*	.03*	–.07*	.31*
Norway	3.43*	–.12	.05*	.02	–.04	.24*
Sweden	2.12*	–.14	.04	.03	.00	.21*
United Kingdom	2.26*	–.05	.14*	.01	–.10	.11
United States	2.91*	.03	.63*	–.04	–.12*	–.07

Note: The OECD demand, OECD export, OECD import, and leftist government coefficients have been multiplied by 10.
* 80 percent confidence interval excludes zero.

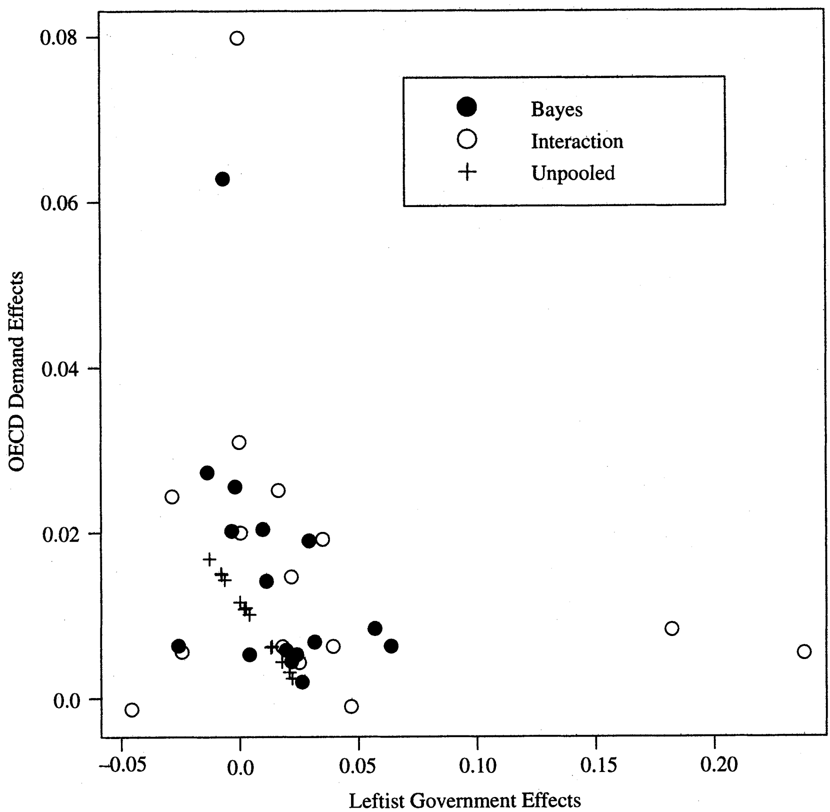
and 1980s. The United States is a large outlier. The unusual size of the U.S. coefficient also may result from the unique influence of American economic growth on demand throughout the OECD area. Here, variation in the time-series coefficients alerts us to the possible endogeneity of OECD demand to U.S. economic growth. The leftist government effects also vary substantially. Large and significant effects are found in Austria and Finland. As the institutional results suggested, we find small demand effects where unions are relatively weak, in Australia, the United Kingdom, and Ireland. Also notice that the model returns leftist government effects for Canada, Japan, and the United States—countries that experienced no leftist government between 1971 and 1984. In these cases the time-series coefficients are almost completely determined by the institutional effects. For instance the U.S. leftist government effect can be approximately calculated from the institutional effects and the U.S. score for labor organization: $-.007 = (.018 \times .86) - .022$. Because data from all countries are used to estimate the institutional effects, the U.S. coefficient is said to be "borrowing strength" from the rest of the sample.

The Bayes shrinkage coefficients can be compared to estimates from two other models. First, the interaction model provides country-level time-series effects that depend linearly on labor organization. For instance, in the notation of Equation 11, the OECD demand effect for country i is $\beta_2 + (\beta_8 \times L_i)$. Second, time-series effects can be obtained from an unpooled analysis that ignores the institutional data and fits models for each country separately. Differences between the three models are illustrated in Figure 1 which plots leftist government and OECD demand effects for each of the sixteen countries in the sample. Coefficients from the interaction model are linearly related and show the least variation. Unpooled estimates, on the other hand, are widely scattered. Because the Bayes estimates are based partly on the institutional data, they are more clustered than the unpooled coefficients which neglect this information. Because they carry an extra component of stochastic variation, however, the Bayes estimates are more dispersed than those from the interaction model. The Bayes estimates can thus be understood as a compromise, admitting more causal heterogeneity than the interaction model but less than the unpooled analysis.

Given the object of describing causal heterogeneity across countries, the unpooled and Bayesian analyses seem preferable to the interaction model in this application. The linear structure of interaction model allows very narrow variation, but the hierarchical and unpooled coefficients are more widely dispersed. Why might we then prefer the Bayes estimates to the unpooled coefficients? First, the Bayes estimates tend to have greater precision than the unpooled coefficients. This can be seen from Figure 2. Each panel of the figure plots the 80 percent confidence intervals for all sixteen countries from Bayesian and unpooled estimation of each time-series effect. If the Bayes and unpooled estimates had equal variance, they would fall on the 45 degree line. Points falling below the 45 degree line indicate the greater precision of the Bayes estimate. The Bayes estimates have lower variance than nearly all the unpooled estimates. In some cases the variance is reduced by as much as 50 percent or more. Variance reduction is largest for the lagged growth effects because these are most imprecisely estimated with the country-level data.

While the value of Bayes as a compromise between the unpooled and interaction models has some intuitive appeal, the superior performance of the hierarchical model can be quantified in out-of-sample forecasting. For this exercise, the models were fit to data from 1971 to 1983, and forecasts were obtained for the growth rate in 1984. Comparisons of forecasts for the unpooled, Bayesian, and interaction models are reported in Table 4. The Bayesian hierarchical model has lower prediction error than either the interaction or the unpooled models. The forecast error of the hierarchical and unpooled analysis are both about 20 percent lower than the error of the interaction model. This suggests that the interaction model underestimates causal

Figure 1. Scatterplots of Country-level Leftist Government and OECD Demand Coefficients from Interaction, Bayesian, and Unpooled Models



heterogeneity. The hierarchical model performs slightly better than the unpooled analysis, suggesting that no important causal complexity is lost by shrinkage in the Bayes estimates.

5. Discussion

Hierarchical models provide substantive and statistical advantages over conventional approaches to pooled data. This paper has focused on models for pooled cross-sectional time-series analysis in comparative research. From a substantive perspective, a key idea of comparative research is that

Figure 2. Eighty Percent Confidence Intervals from Bayes and Unpooled Least Squares Coefficients for (a) Average Growth, (b) Lagged Growth, (c) OECD Demand, (d) OECD Exports, (e) OECD Imports, (f) Leftist Government

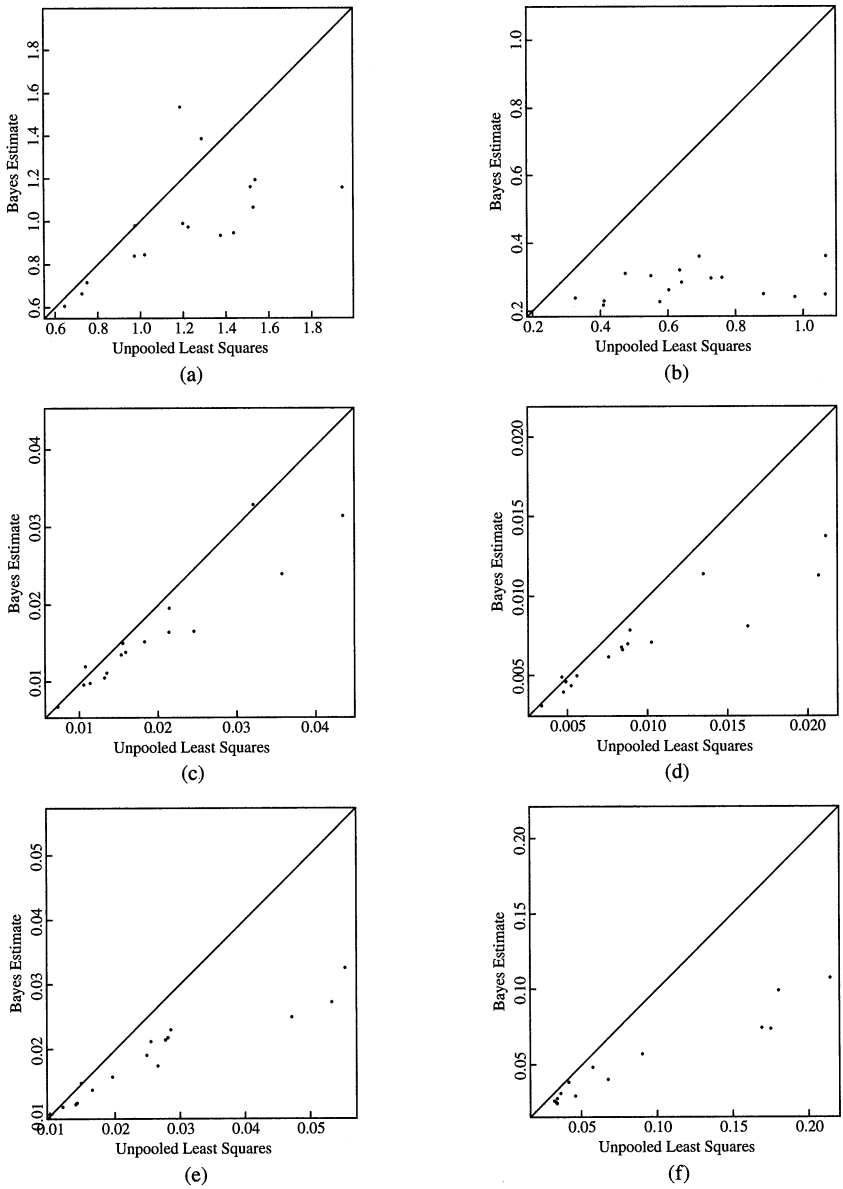


Table 4. Out-of-Sample Squared Prediction Error for Interaction, Hierarchical and Unpooled Models

	Model		
	Unpooled	Interaction	Hierarchical
Australia	16.4	8.4	8.5
Austria	0.0	3.4	11.0
Belgium	20.6	23.0	16.7
Canada	3.8	1.0	2.9
Denmark	0.2	0.0	0.3
Finland	0.2	0.5	0.8
France	4.2	11.4	6.4
Germany	3.3	7.2	4.3
Ireland	8.7	11.8	0.9
Italy	13.2	5.7	12.7
Japan	2.0	0.0	1.8
Netherlands	6.6	23.7	9.6
Norway	11.1	11.5	9.9
Sweden	1.2	0.2	0.6
United Kingdom	5.2	7.3	8.4
United States	0.2	6.1	0.7
Total	97.0	121.2	95.0

causal processes vary across countries. The fundamental problem of comparative research is contextual explanation where differences in causal processes within countries are related to characteristics that vary across countries. This contextual idea of comparative analysis is especially relevant to comparative political economy. In this area, cross-national variation in economic relationships originates with enduring institutional differences. Hierarchical models provide a powerful tool for examining these sorts of contextual effects. In contrast to standard models for pooled cross-sectional time-series data, hierarchical models allow for causal heterogeneity across countries. Time-series coefficients are estimated for each country, and these are specified to depend on time-invariant institutional conditions. Model dynamics will often be substantively important, and these can be included in the structural part of the model.

Hierarchical models also provide several important statistical advantages. The models yield estimates of institutional effects that take account of uncertainty about time-series effects. This uncertainty is neglected in naive regressions on time-series coefficients. Shrinkage estimates of time-series coefficients based on the pooled data are statistically superior to least squares estimates based on unpooled data. They are more accurate and yield more accurate out-of-sample predictions. Hierarchical models also offer a more realistic assessment of uncertainty in comparative data than interaction models. Interaction models assume that time-series effects depend deterministically on institutional contexts, ruling out unexplained variation in time-series coefficients. Lastly, hierarchical models also make a conceptual contribution to the general principles of comparative analysis. The idea of exchangeability which underpins the hierarchical model provides a precise definition of "comparability" in comparative research. In essence, we can regard two countries as comparable when we are indifferent about the residual variation that distinguishes them. This will be so when we can view the countries as realizations of a single underlying model.

These ideas were illustrated in a reanalysis of the OECD economic growth data of Alvarez, Garrett, and Lange (1991). The analysis showed that enduring cross-national variation in labor organization influenced not just the impact of leftist government on economic growth, but the institutionalized power of labor movements also influenced the impact of surrounding economic conditions. Similar to Alvarez, Garrett, and Lange (1991) and Beck et al. (1993), the hierarchical model provided evidence that leftist government only positively affected economic growth in countries with encompassing and centralized unions. In addition, the hierarchical model provided new evidence that strong unions movements could insulate countries from the negative effects of declining international demand.

The hierarchical model presented here can be generalized in several ways. Additional layers could be added to the model hierarchy. We might have time-series data for a large number of electoral districts for many countries. We could then estimate separate equations for each district, country-level equations to explain variation across districts, and finally cross-national equations to explain variation in the country-level coefficients. This may sound impractical but simpler models could add a single random effect for each district allowing for unobserved heterogeneity across districts. The hierarchical modelling approach could also be applied to other kinds of data. I've focused here on the pooled cross-sectional time-series design but the same ideas could also be applied to survey data from many countries or survey data from a single country at many points in time. Such data are not uncommon in political research, but pooled hierarchical modelling has rarely been applied. Finally, hierarchical models could be extended beyond the linear case to include generalized linear models which feature logistic and Poisson regression as special cases. Accurate estimate of these models can be obtained with the Gibbs sampler (Rodríguez and Goldman 1995). Applications of these models are already proliferating in other fields (see the review of DiPrete and Forristal 1994). Like the hierarchical model presented here, these extensions provide a flexible and statistically powerful class of specifications for examining the pooled data of comparative politics. More highly parameterized than conventional models, hierarchical models allow the investigation of complex causal stories while providing a realistic assessment of uncertainty arising under a comparative research design.

Manuscript submitted 3 May 1997.
Final manuscript received 9 December 1997.

REFERENCES

Alvarez, R. Michael, Geoffrey Garrett, and Peter Lange. 1991. "Government Partisanship, Labor Organization, and Macroeconomic Performance." *American Political Science Review* 85: 539–56.

Bartels, Larry. 1996. "Pooling Disparate Observations." *American Journal of Political Science 40*: 905–42.

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What to Do (and Not to Do) with Time-Series Cross-Section Data." *American Political Science Review* 89:634–47.

Beck, Nathaniel, and Jonathan N. Katz. 1996a. "Nuisance vs. Substance: Specifying and Estimating Time-Series–Cross-section Models." *Political Analysis* 6.

Beck, Nathaniel and Jonathan N. Katz. 1996b. "Lumpers and Splitters United: The Random Coefficients Model." Presented at the annual meeting of the Political Methodology Group. Ann Arbor, Michigan.

Beck, Nathaniel, Jonathan Katz, R. Michael Alvarez, Geoffrey Garrett, and Peter Lange. 1993. "Government Partisanship, Labor Organization, and Macroeconomic Performance: A Corrigendum." *American Political Science Review* 87:945–948.

Bruno, Michael, and Jeffrey Sachs. 1985. *The Economics of Worldwide Stagflation.* Oxford: Blackwell.

Bryk, Anthony S., and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park, Calif.: Sage.

Casella, George, and Edward I. George. 1992. "Explaining the Gibbs Sampler." *American Statistician* 46:167–174.

De Leeuw, Jan, and Ida Kreft. 1986. "Random Coefficient Models for Multilevel Analysis." *Journal of Educational Statistics* 11:57–85.

DiPrete, Thomas A., and Jerry D. Forristal. 1994. "Multilevel Models: Methods and Substance." *Annual Review of Sociology* 20:331–57.

Draper, David, James S. Hodges, Colin L. Mallows, and Daryl Pregibon. 1993. "Exchangeability and Data Analysis." *Journal of the Royal Statistical Society A* 156:9–37.

Efron, Bradley, and Robert J. Tibshirani. 1993. *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Erbring, Lutz, and Alice A. Young. 1980. "Individuals and Social Structure: Contextual Effects as Endogenous Feedback." In *Aggregate Data: Analysis and Interpretation,* eds. Edgar Borgatta and David Jackson. Beverly Hills: Sage.

Gelman, Andrew B., John B. Carlin, Hal S. Stern, and Donald B. Rubin. 1995. *Bayesian Data Analysis.* London: Chapman and Hall.

Griffin, Larry J., Philip J. O'Connell, and Holly J. McCammon. 1989. "National Variation in the Context of Struggle: Postwar Class Conflict and Market Distribution in the Capitalist Democracies." *Canadian Review of Sociology and Anthropology* 26:37–68.

Hicks, Alexander M. 1994a. "The Social Democratic Corporatist Model of Economic Performance in the Short- and Medium-Run Perspective." In *The Comparative Political Economy of the Welfare State,* ed. Thomas Janoski and Alexander Hicks. New York: Cambridge University Press.

Hicks, Alexander M. 1994b. "Introduction to Pooling." In *The Comparative Political Economy of the Welfare State,* ed. Thomas Janoski and Alexander Hicks. New York: Cambridge University Press.

Huber, Evelyne, Charles C. Ragin, and John D. Stephens. 1993. "Social Democracy, Christian Democracy, Constitutional Structure, and the Welfare States." *American Journal of Sociology* 99:711–49.

Iversen, Gudmund R. 1991. *Contextual Analysis.* Newbury Park, Calif.: Sage.

Katzenstein, Peter. 1985. *Small States in World Markets.* Ithaca, N.Y.: Cornell University Press.

Korpi, Walter. 1989. "Power, Politics and State Autonomy in the Development of Social Citizenship: Social Right During Sickness in Eighteen OECD Countries Since 1930." *American Sociological Review* 54:309–28.

Kreft, Ita G.G., Jan de Leeuw, and Rien van der Leeden. 1994. "Review of Five Multilevel Analysis Programs: BMDP-5V, GENMOD, HLM, ML3, VARCL." *American Statistician* 48:324–35.

Lange, Peter and Geoffrey Garret. 1985. "The Politics of Growth." *Journal of Politics* 47:792–827.

Lieberson, Stanley. 1992. "Small *N*'s and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." In *What is a Case? Exploring the Foundations of Social Inquiry,* eds. Charles Ragin and Howard S. Becker. New York: Cambridge University Press.

Lindley, D.V. and A.F.M. Smith. 1972. "Bayes Estimates for the Linear Model." *Journal of the Royal Statistical Society* Series B 34:1–41.

Mason, William M., G. Y. Wong, and Barbara Entwistle. 1983. In *Sociological Methodology 1983–1984,* ed. S. Leinhardt. Oxford: Blackwell.

Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry.* New York: Wiley-Interscience.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies.* Berkeley: University of California Press.

Robinson, G.K. 1991. "That BLUP is a Good Thing: The Estimation of Random Effects." *Statistical Science* 6:15–51.

Rodríguez, Germán, and Noreen Goldman. 1995. "An Assessment of Estimation Procedures for Multilevel Models with Binary Responses." *Journal of the Royal Statistical Society Series A* 158:73–89.

Scharpf, Fritz. 1991. *Crisis and Choice in European Social Democracy.* Ithaca, N.Y.: Cornell University Press.

Schumpeter, Joseph A. 1954. *History of Economic Analysis.* New York: Oxford University Press.

Solow, Robert M. 1986. "Unemployment: Getting the Questions Right." *Economica* 53:S23–S34.

Spielgelhalter, David, Andrew Thomas, Nicky Best, and Wally Gilks. 1996. *BUGS 0.5: Bayesian Inference Using Gibbs Sampling, Manual (version ii).* MRC Biostatistics Unit, Institute of Public Health. Cambridge.

Stimson, James A. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29:915–47.

Swank, Duane. 1992. "Politics and the Structural Dependence of the State in Democratic Capitalist Nations." *American Political Science Review* 86:38–54.

Tanner Martin. 1993. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions.* 2nd ed. New York: Springer-Verlag.

Thelen, Kathleen, and Sven Steinmo. 1992. "Historical Institutionalism in Comparative Politics." In *Structuring Politics: Historical Institutionalism in Comparative Analysis,* eds. Sven Steinmo, Kathleen Thelen, and Frank Longstreth. New York: Cambridge University Press.

Tilly, Charles. 1984. *Big Structures, Large Processes, Huge Comparisons.* New York: Russell Sage Foundation.

Western, Bruce. 1997a. "Explaining the OECD Wage Slowdown: Recession or Labor Decline?" Presented at the Conference on Unemployment in Southern Europe in Comparative Perspective. Princeton University, Princeton NJ. November.

Western, Bruce. 1997b. *Between Class and Market: Postwar Unionization in the Capitalist Democracies.* Princeton: Princeton University Press.

Western, Bruce, and Simon Jackman. 1994. "Bayesian Inference for Comparative Research." *American Political Science Review* 88:412–23.