<u>Modeling Spatial Relationships in International and Comparative Political Economy</u>: An Application to Globalization and Capital Taxation in Developed Democracies

Robert J. Franzese, Jr.
Associate Professor of Political Science
The University of Michigan, Ann Arbor
franzese@umich.edu

Jude C. Hays
Assistant Professor of Political Science
The University of Michigan, Ann Arbor
jchays@umich.edu

28 March 2003

<u>Abstract</u>: Social scientists have long recognized that time-series-cross-section (*TSCS*) data typically correlate across time and space. Today, standard social-science practice is to model dynamics (*i.e.*, temporal dependence) directly, typically with lags of the dependent variable, but to address spatial dependence solely by applying panel-corrected (robust) standard-errors (*PCSE*), thereby treating spatial dependence as a *nuisance* (Beck and Katz's 1996 terminology). However, regardless of the analyst's substantive interest in these relationships, direct modeling of spatial dependence (plus *PCSE* perhaps) enhances efficiency and is required for unbiased and consistent estimates of the coefficients on non-spatial regressors and any associated hypothesis tests or confidence intervals. Whereas analysts less interested in spatial relations *per se* will understandably wish to employ simple proxies for more complicated diffusion processes, those directly interested will prefer more sophisticated modeling techniques for estimating the dyadic patterns of diffusion in their data. Our broad project uses Monte Carlo experiments to evaluate the performance of several simple and sophisticated estimators under three important types of spatial correlation (the more-complex being spatial analogues to estimators devised for dynamic panel models: *e.g.*, Hsiao 1986; Baltagi 1995) and develops a set of techniques and guidelines to help analysts diagnose, characterize, and gauge various sorts of spatial correlation, and to choose and interpret appropriate estimators based on their objectives. This paper leverages globalization and capital taxation as substantive venue to start such methodological explorations. Many academic and casual observers argue that the dramatic post-1970s rise in international capital mobility and the steadily upward postwar trend in trade integration, by sharpening capital's threat against domestic governments to flee purportedly excessive and inefficient taxation, has forced and will continue to force welfare-/tax-state retrenchment and tax-burden shifts away from more-mobile capital (especially financial capital) and toward less-mobile labor (especially manual labor). Several important recent studies of the international and comparative political economy of policy change over this period challenge such claims. We offer a preliminary comparison of several such arguments, using specifications that reflect the spatial relationships central to such diffusion processes.

Modeling Spatial Relationships in International and Comparative Political Economy: An Application to Globalization and Capital Taxation in Developed Democracies

Abstract: Social scientists have long recognized that time-series-cross-section (*TSCS*) data typically correlate across time and space. Today, standard social-science practice is to model dynamics (*i.e.*, temporal dependence) directly, typically with lags of the dependent variable, but to address spatial dependence solely by applying panel-corrected (robust) standard-errors (*PCSE*), thereby treating spatial dependence as a *nuisance* (Beck and Katz's 1996 terminology). However, regardless of the analyst's substantive interest in these relationships, direct modeling of spatial dependence (plus *PCSE* perhaps) enhances efficiency and is required for unbiased and consistent estimates of the coefficients on non-spatial regressors and any associated hypothesis tests or confidence intervals. Whereas analysts less interested in spatial relations *per se* will understandably wish to employ simple proxies for more complicated diffusion processes, those directly interested will prefer more sophisticated modeling techniques for estimating the dyadic patterns of diffusion in their data. Our broad project uses Monte Carlo experiments to evaluate the performance of several simple and sophisticated estimators under three important types of spatial correlation (the more-complex being spatial analogues to estimators devised for dynamic panel models: *e.g.*, Hsiao 1986; Baltagi 1995) and develops a set of techniques and guidelines to help analysts diagnose, characterize, and gauge various sorts of spatial correlation, and to choose and interpret appropriate estimators based on their objectives. This paper leverages globalization and capital taxation as substantive venue to start such methodological explorations. Many academic and casual observers argue that the dramatic post-1970s rise in international capital mobility and the steadily upward postwar trend in trade integration, by sharpening capital's threat against domestic governments to flee purportedly excessive and inefficient taxation, has forced and will continue to force welfare-/tax-state retrenchment and tax-burden shifts away from more-mobile capital (especially financial capital) and toward less-mobile labor (especially manual labor). Several important recent studies of the international and comparative political economy of policy change over this period challenge such claims. We offer a preliminary comparison of several such arguments, using specifications that reflect the spatial relationships central to such diffusion processes.

I. Introduction

Recognizing that observations in time-series-cross-section (*TSCS*) datasets usually correlate across time and space, social scientists now typically model dynamics (*i.e.*, temporal dependence) directly, often with lags of the dependent variable, but address spatial dependence solely by applying panel-corrected (robust) standard-errors (*PCSE*), thereby treating spatial dependence as a *nuisance* (Beck and Katz's 1996 terminology). As we explain below, however, direct modeling of spatial dependence (plus robust standard-errors perhaps) is superior, regardless of the analyst's substantive interest in these relationships. Directly modeling spatial dependence always enhances efficiency and, moreover, is often critical to obtain unbiased and consistent estimates of the coefficients on non-spatial regressors and all associated hypothesis tests and confidence intervals. Understandably, analysts less interested in spatial relationships *per se* would prefer simple proxies for complicated diffusion processes. Those directly interested, conversely, will prefer more sophisticated modeling techniques for estimating the dyadic patterns of diffusion in their data. In this project, we conduct multiple Monte Carlo experiments to evaluate the performance of several simple and sophisticated estimators under three important types of spatial correlation (the more-complex methods being spatial analogues to estimators devised for dynamic panel models: *e.g.*, Hsiao 1986; Baltagi 1995). We also develop a set of techniques and guidelines to help analysts diagnose, characterize, and gauge spatial correlation and to choose and interpret appropriate estimators based on their objectives.

We thus address both social-science researchers directly interested in spatial relationships (*substance*) and those primarily concerned to make optimal inferences regarding other substantive relationships given spatially dependent data (*nuisance*). Building from analogies to similar, better-explored issues arising from temporal dependence, and through analytic derivation and Monte Carlo experimentation, we will: (1) detail conditions under which failing to model spatial dependence or relegating its role to standard-error adjustment appreciably biases other coefficient estimates or mostly induces *mere* inefficiency, exploring the magnitudes of these biases and inefficiencies under varying degrees and natures of spatial dependence; (2) distinguish *spatial diffusion* from *spatially correlated responses to omitted spatially-correlated factors* conceptually and evaluate alternative approaches to distinguishing such spatial relations empirically; (3) develop and explore the properties of several parametric, semi-, and non-parametric methods of testing for spatially correlated disturbances, with and without spatial lags in the model; (4) compare the properties of simple proxies for full models of the true spatial-diffusion processes or common omitted factors—*e.g.*, spatial dummies or symmetric spatial-lags that average other cross-section units' dependent variables each time-period—to each other with and without *PCSE*, to *PCSE* alone, and to various methods of estimating the full model; and (5) explore the properties of leveraging spatial dependence to aid identification of endogenous systems by using spatial lags as *quasi-instruments* (Bartels 1991) under differing magnitudes of the relevant spatial correlations and causal arrows: $\mathbf{X}\leftrightarrow\mathbf{Y}$, $\mathbf{X_i}\leftrightarrow\mathbf{X_j}$, and $\mathbf{Y_i}\leftrightarrow\mathbf{Y_j}$ (or $\boldsymbol{\varepsilon}_\mathbf{i}\leftrightarrow\boldsymbol{\varepsilon}_\mathbf{j}$). As we aim to assist analysts of varying statistical sophistication about, and substantive interest in, spatial relationships to develop useful intuitions and techniques to diagnose, gauge, and characterize spatial correlation and to choose and interpret appropriate estimators for their objectives, we will also (6) create and disseminate statistical-software algorithms to implement all of our suggested techniques and, where possible and productive, pedagogical modules to help teach them.[1]

This paper studies globalization and capital taxation as a substantive venue from which to start such methodological explorations. Many academic and casual observers argue that the dramatic post-1972 rise in global capital mobility and the steady postwar rise in trade integration sharpen capital's threat against domestic governments to flee "excessive and inefficient" taxation and public policies. This, the standard view holds, has forced and will continue to force welfare/tax-state retrenchment and tax-burden shifts from more-mobile capital (especially financial) toward less-mobile labor (especially skilled-manual). Recently, several important studies of the international and comparative political economy of tax/welfare policy over this era challenge such claims on at least four distinct bases: Garrett (1998) argues certain combinations of left government with social-welfare, active-labor-market, coordinated-bargaining, and related policies can be as or more efficient than neoliberal state-minimalism and conservative government and, therefore, that capital will not

[1] Beck and Katz, who introduced the extremely useful and now almost universally employed *PCSE*s, have a current project that likewise addresses some gaps in political methodology regarding spatial dependence. Although they intend to explore some related applications of spatial lags to model spatial dependence, our foci differ from and complement theirs in that and in other regards. They propose to consider (a) non-geographic analogues to spatial-lag models of contemporaneous correlation, (b) modeling approaches to spatial parameter-heterogeneity, and models of spatial dependence in (c) nominal and in (d) non-stationary data. Complementarily, we propose to explore (a) the conditions that determine whether and to what degree failing to model spatial dependence directly biases other coefficient estimates or *merely* induces inefficiency, (b) approaches to distinguishing spatial diffusion from spatially correlated responses conceptually and empirically, (c) tests for and gauges of spatially correlated disturbances, with and without modeled spatial lags, (d) comparison of simple spatial-indicator and spatial-lag proxies to each other, to *PCSE* alone, and to fuller, more-complex spatial-dependence models, and (e) leveraging of spatial dependence to aid identification of endogenous systems by using spatial lags as *quasi-instruments* (Bartels 1991).

flee such efficient combinations; Boix (1998) argues public (human and physical) capital-investment strategies comprise an alternative to neoliberal minimalism that is sufficiently economically efficient to attract and retain capital and politically effective to maintain left electoral-competitiveness; Hall and Soskice (2001) argue complex national networks of political-economic institutions confer *comparative* advantages in differing productive activities, which, as Mosher and Franzese elaborate (2002), implies capital mobility and trade integration (*if international tax-competition remains sufficiently muted*) would spur institutional and policy specialization—here, cross-national welfare/tax-state variation—rather than convergence or global retrenchment; and Swank (2002) argues that the institutional structure of the polity and of the welfare/tax system itself shape domestic policy responses to capital (and trade) integration. We review several such arguments and offer a preliminary evaluation, specifying empirical models that, unlike these and other previous efforts, embody the spatial relationships central to such diffusion processes.

Part II of the paper discusses our conceptualization of the types of spatial relationships that may characterize data in international and comparative political economy contexts and the issues they raise for empirical evaluation of social-political-economic theories in such contexts. Part III illustrates some of these problems in data exhibiting each of three crucial types of spatial dependence: spatially correlated stochastic components (*i.e.*, error terms) with exogenous and orthogonal regressors, with exogenous but spatially correlated regressors, and with endogenous and correlated regressors. Part IV elaborates and clarifies the methodological analyses planned for the broader project. Part V surveys the theoretical propositions regarding financial-capital mobility and capital-taxation convergence, and Part VI constructs and conducts our empirical model and analysis. Part VII concludes with some discussion of next steps.

II. Conceptualizing Spatial Relationships

Much political-science research uses time-series-cross-section data: datasets that contain observations on several cross-sectional units over multiple time-periods. The units are typically interdependent, and this dependence, which may or may not relate to geographic proximity, is termed *spatial*. The spatially correlated units will often correlate *contemporaneously*. (Note that *contemporaneousness* depends upon the data's level of temporal aggregation. Spatial processes that take weeks or months appear instantaneous in annual data.) While political scientists widely recognize this interdependence, the difficulties that spatial correlation can create for empirical research have received little consideration beyond Beck and Katz's seminal work. We seek to redress key aspects of these gaps in current political methodology.

We begin conceptually, noting two distinct sources of spatial correlation: *spatial diffusion* across units and *common* (or *correlated*) *shocks* to all units. In the former process, an internal change within one unit affects other units. Common shocks, contrarily, originate from an external source and affect multiple units simultaneously. Outcomes in several oil-dependent economies, *e.g.*, may correlate spatially because each experiences the same external price-shocks. To develop these ideas formally, start with a basic model:

$$1. \qquad \mathbf{y} = \mathbf{X}\beta + \varepsilon$$

where $\mathbf{y}$ is a *TxN* matrix of *T* outcomes (1 *per* time period) for *N* units (1 *per* cross-sectional/spatial unit), $\mathbf{X}$ is a *TxNK* matrix of values on *K* independent variables, $\boldsymbol{\varepsilon}$ is a *TxN* matrix of spatially correlated errors, and $\boldsymbol{\beta}$ is an *NKxN* matrix of coefficients. To simplify exposition, we assume henceforth that $K=1$ and that the coefficients $\boldsymbol{\beta}$ are identical in each spatial unit. Under these assumptions, $\boldsymbol{\beta}$ simplifies to an *NxN* matrix with common diagonal element β and zeros off-diagonal. (More generally, $\boldsymbol{\beta}=\mathbf{I}\otimes\boldsymbol{\Sigma}$, where $\mathbf{I}$ is an *NxN* identity matrix and $\boldsymbol{\Sigma}$ is a *Kx1* vector of coefficients.)

Decomposing $\boldsymbol{\varepsilon}$ into component parts, we have

$$2. \qquad \mathbf{y} = \mathbf{SW}\rho + \mathbf{X}\beta + \eta + \upsilon$$

where $\mathbf{W}$ is an NxN spatial diffusion matrix, which contains zeros along the diagonal ($w_{ii}=0$) and some non-zero elements off the diagonal; the cell w_{ij} measures the degree of diffusion from unit i to unit j and w_{ji} measures the degree of diffusion from unit j to unit i. Note that $\mathbf{W}$, which fully describes the pattern of spatial diffusion, has up to $N \cdot (N-1)$ unique elements. Thus, the number of spatial-diffusion terms to estimate grows much faster than N, which, except for exceptionally large T relative to N, underscores the magnitude of the challenge spatial diffusion presents for empirical researchers, and explains why those directly interested in such spatial relationships require some model (*i.e.*, theoretically reduced parameterization) of the diffusion process, and why those less-substantively interested require some simplified proxy for them. The matrix $\boldsymbol{\eta}$ is a TxN matrix with a common row element, representing the spatially common component of the shock in each time period; and $\boldsymbol{\upsilon}$ is a TxN matrix with independent elements, *i.e.*, each element represents the component of the shock unique to each spatial unit in each time period. The TxN matrix $\mathbf{S}$ is the sum of $\mathbf{X}$, $\boldsymbol{\eta}$, and $\boldsymbol{\upsilon}$, so that $\mathbf{SW}$ is the *spatial lag* (see Anselin 1988), each element of which is a diffusion-matrix-weighted average of the independent variables and shocks elsewhere in the cross-section that time period. Similar to $\boldsymbol{\beta}$, the NxN matrix $\boldsymbol{\rho}$ has zeros off the diagonal and a common diagonal element ρ, giving the coefficient on the spatial lag. With these definitions, we can write the equation for each element of the outcome matrix as

$$3. \qquad y_{n,t} = \rho \mathbf{s}_t \mathbf{w}_t + \beta x_{n,t} + \eta_t + \upsilon_{n,t}$$

where $\mathbf{s_t}$ ($1xN$) is the sum of the t^{th} rows from $\mathbf{X}$, $\boldsymbol{\eta}$, and $\boldsymbol{\upsilon}$, and $\mathbf{w_n}$ ($Nx1$) is the n^{th} column from $\mathbf{W}$. Note the first and penultimate terms reflect what we call *spatial diffusion* and *common shocks* respectively.

Most political scientists view such spatial dependence as *nuisance* (Beck and Katz 1996), relegating it to a mere standard-error-adjustment role *via* panel-corrected (robust) standard-errors (*PCSE*). That so few political scientists have modeled the spatial relationships in their data directly is surprising (cf. Ward and O'Loughlin 2002) because these spatial relationships not only map the spatially correlated effects of shocks (penultimate term) but also the cumulative reverberation of shocks and of modeled changes *via* diffusion (first term). As demonstrated here, best practice is always to model such spatial dependence, regardless of one's substantive interest in these relationships because this approach adds efficiency and, under many circumstances, is crucial to avoid sizable bias and inconsistency in the estimates of coefficients on even the non-spatial regressors (and any hypothesis tests and confidence intervals regarding them).

Understandably, analysts less interested in spatial relationships *per se* will prefer simple proxies for such complex processes, whereas those more-directly interested will prefer more-sophisticated modeling techniques to estimate the substantively rich patterns of diffusion in their data. Accordingly, we explore three methods of rising complexity to address spatial correlation. One simple approach is to include time-period dummies. In theory, this approach assumes that the spatial correlation arises from external *common shocks* alone, but, in practice, the dummies will also partially proxy for any spatial diffusion present. A slightly less simplistic method might include average values of the dependent variable in the *N-1* other cross-sectional units for each time period (see, *e.g.*, Franzese 1999, 2002, 2003b). Theoretically, this approach assumes a symmetric spatial-diffusion process (*i.e.*, each unit affects all other units and all other units affect it equally) without common shocks, but again, in practice, it will proxy partially for both asymmetrical diffusion and common shocks. Thus, the estimated coefficients on either of these simple spatial proxies will have ambiguous interpretation in practice, although this may be less problematic for analysts lesscentrally interested in interpreting the spatial dependence. More thorough approaches would estimate the patterns and strength of spatial diffusion in one's data, for which, again, relatively simpler or more-complicated estimation techniques exist (*e.g.*, two-stage or FIML estimation).

Accordingly, one part of our project will evaluate how well, under varying independent-variable and stochastic-term conditions, these and other such simple or complex approaches to spatial dependence can distinguish spatial diffusion from common shocks, and how well they can estimate the spatial and non-spatial regressors' coefficients and standard errors, relative to each other with and without *PCSE* and to estimating *PCSE* alone. This evaluation favors direct modeling of spatial dependence, simpler or more-complex depending on the degree and pattern of those correlations, so researchers will logically need tools to assess the spatial correlation in their data and, ultimately, in their estimated residuals. Most critically, we note that, as with temporal lags, spatial lags induce bias if and to the degree that stochastic components retain correlation controlling for the model of the systematic component. Accordingly, our project will next develop and explore the properties of several parametric, non-, and semi-parametric tests for and gauges of spatial correlation in the presence or absence of spatial lags in the model.[2] Then, we explore the properties of leveraging spatial correlation to help identify endogenous systems of equations. In political economy, *e.g.*, many researchers (*e.g.*, Alvarez at al. 1991; Beck et al. 1993, Franzese 1999, 2002, 2003b) include economic conditions abroad as controls or substantively central factors in their analyses, assuming them exogenous. Insofar as exogeneity actually holds, these factors offer potentially useful instruments to identify endogenous systems. Unemployment abroad, *e.g.*, might instrument for domestic unemployment to estimate its effect on government spending, which latter may partly cause the former, implying endogeneity and the need to identify by instrumentation or some other

[2] We have begun one such exploration below.

means. However, if outcomes abroad affect domestic conditions, one criterion for a useful instrument, then quite likely domestic conditions affect outcomes abroad, violating the other condition for a perfect instrument. Thus, researchers attempting to leverage spatial correlation to identify endogenous systems, or even simply employing external conditions as an assumed-exogenous factor explaining domestic outcomes—and possible applications to virtually any study in *TSCS* data—must recognize that such external conditions are partly endogenous and so only *quasi-instruments* (Bartels 1991). Thus, we will explore the properties of leveraging spatial dependence to aid identification of endogenous systems by using spatial lags as *quasi-instruments* (Bartels 1991) under differing magnitudes of the relevant spatial correlations and causal arrows: $\mathbf{X}\leftrightarrow\mathbf{Y}$, $\mathbf{X_i}\leftrightarrow\mathbf{X_j}$, and $\mathbf{Y_i}\leftrightarrow\mathbf{Y_j}$ (or, equivalently, $\boldsymbol{\varepsilon_i}\leftrightarrow\boldsymbol{\varepsilon_j}$). Finally, we intend also to create and disseminate statistical-software algorithms and pedagogical modules as possible and useful for our suggested techniques.

In these explorations, we will consider at least three types of spatial dependence: 1) spatially correlated disturbances with orthogonal, exogenous regressors, 2) spatially correlated disturbances with spatially correlated, exogenous regressors, and 3) orthogonal disturbances with spatially correlated and endogenous regressors. We introduce these three cases in the next part, describing the methodological problems that arise in each and highlighting the questions for which researchers who work with *TSCS* data, both those directly interested in spatial relations and those not, need answers. To illustrate, we offer the results of one-shot simulations for each case. Obviously, these simulations and the discussion below are preliminary, one central task of this project being, of course, to develop these one-shot simulations into a full set of Monte Carlo experiments. In the following Part III, we review our plan of research for the methodological project before undertaking the substantive application and concluding in the final three parts.

III. Three Forms of Spatial Dependence and the Associated Empirical Challenges

1. Spatially Correlated Disturbances with Exogenous and Orthogonal Regressors

We start with the case of spatially correlated errors and spatially orthogonal regressors, which is represented by the following set of variance-covariance matrices:

$$4. \qquad \mathrm{var}(\varepsilon) = \sigma^2\mathbf{C} \;\; \textit{and} \;\; \mathrm{var}(\mathbf{X}) = \psi^2\mathbf{I}$$

The parameters σ^2 and ψ^2 are scalars. The matrix **C** is *NxN* with ones on the diagonal and non-zero off-diagonal elements. The variance-covariance matrix **var[ε]** is, therefore, an *NxN* matrix with σ^2 on the diagonal and non-zeros off it. These off-diagonals are a function of *common shocks*. The matrix ***I*** is an *NxN* identity matrix. The variance-covariance matrix **var[X]** is *NxN* with ψ^2 on the diagonal and zeros off it.

a. Efficiency in Estimating β (Table 1, Analyses II-III)

One reason analysts might wish to model spatial relationships in their *TSCS* data is to enhance the estimation precision of their non-spatial regressors, aiming to *absorb* some residual variance to reduce the standard errors of $\hat{\beta}$. To check this intuition and gain a sense of the magnitude of these efficiency gains, Table 1, Analyses I-III show the results of a one-shot simulation. The true model that generated the simulation dataset of 20 cross-section units over 30 time periods is $y_{n,t} = .2\mathbf{s}_t\mathbf{w}_n + .8x_{n,t} + \eta_t + \upsilon_{n,t}$, with each $x_{n,t}$, η_t, and $\upsilon_{n,t}$ a *N(0,1)* random draw and **W** this, arbitrary, asymmetrical spatial-diffusion matrix:

5. $$\mathbf{W}=\begin{pmatrix}
0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 \\
0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 \\
0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 \\
0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 \\
0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 \\
0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 & 0.6 \\
0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 & 0.7 \\
0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.8 \\
0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 \\
0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

As a benchmark against which to gauge the efficiency gains, we regress **Y** on a constant and **X**. The estimated coefficient on **X** will be inefficient but unbiased, because the omitted spatially correlated factors do not correlate with the included independent (*ergo* spatially uncorrelated) $x_{n,t}$ draws. Table 1 (Column 1) reports these results. The estimated standard error on $\hat{\beta}$, which is biased, is almost 0.10.

We consider two simple ways to model the spatial correlation. One approach includes the average value of the dependent variable in other cross-section units each period, which we denote $\bar{y}_{\sim n,t}$, on the right-hand-side. $\bar{y}_{\sim n,t}$ essentially proxies for η_t and $\mathbf{s_t w_n}$ here. The other simple approach adds time-period dummies. Severe multicolinearity will often prohibit estimating models with both $\bar{y}_{\sim n,t}$ and time dummies, underscoring the importance of determining conditions under which one outperforms the other. With **X** orthogonal to the spatial lag and common shocks, we expect either strategy to estimate $\hat{\beta}$ without severe bias; which strategy will produce more-efficient $\hat{\beta}$ estimates is unclear *a priori*. ($\hat{\beta}$ estimates will likely exhibit some bias with either $\bar{y}_{\sim n,t}$ or time dummies because both proxies tend to be endogenous. Analogously to including lagged dependent variables when error terms retain autocorrelation, spatial-proxy estimates will be biased, thereby biasing other coefficient estimates.) The efficiency of time dummies relative to $\bar{y}_{\sim n,t}$ will presumably depend on the share of **Y**'s total variation explained by common shocks relative to the number of time periods. At small ratios, efficiency costs from lost degrees of freedom will outweigh the dummies explanatory-power advantage, so $\bar{y}_{\sim n,t}$ will dominate, and *vice versa* at large ratios.

In the one-shot simulation (Table 1), the time dummies produce a slightly (and probably negligibly) more precise $\hat{\beta}$ estimate (Analysis II *vs.* Analysis III), but this will not always hold. Both dominate the simple regression of **Y** on **X**, cutting the standard error by more than half. Note that we have not yet explored how accurately this nominal standard-error reduction reflects any true increased precision, as that would require Monte Carlo experimentation. Note also that one can interpret the estimated coefficient on $\bar{y}_{\sim n,t}$ in Analysis II as an estimate neither of ρ nor of some average common-shock. The estimated coefficients on the period dummies likewise lack simple interpretation.

Table 1: Spatially Correlated Disturbances with Orthogonal Regressors (Simulated Data)

(True Model: $y_{n,t} = .2 * \mathbf{s_t w_n} + .8 * x_{n,t} + \eta_t + \upsilon_{n,t}$)

Variable	Analysis I	Analysis II	Analysis III	Analysis IV	Analysis V
X	0.62645**	0.73229**	0.73618**	0.78320**	0.75976**
	(.09665)	(.04264)	(.04216)	(.04682)	(.04361)
$\bar{y}_{\sim n,t}$		0.99918**			
		(.02104)			
SW				0.38972**	0.08439**
(Spatial Lag)				(.01038)	(.02622)
Fixed Period Effects	No	No	Yes	No	Yes
No. of Observations	600	600	600	600	600

**p-value < .01, *p-value < .05

b. Estimating the Coefficient on the Spatial Lag (Analyses IV-V)

Seeking unbiased and efficient estimates of spatial-lag coefficients, ρ, as a substantively interesting *diffusion process* complicates matters. Methodologically, the problem surrounds difficulty distinguishing *spatial diffusion* from *common shocks*, especially in small samples, and these difficulties arise regardless of whether analysts' models seek to make this distinction. We assume for now that the analyst knows and so need not estimate the spatial-diffusion matrix, **W**, thus isolating the problem distinguishing *common shocks* from *spatial diffusion*. Table 1, Analyses IV-V reports the results. If one ignores the *common shocks*, these omitted fixed period effects will correlate positively with the true spatial lag, biasing the $\hat{\rho}$ estimate upward. In the one-shot simulation (Analysis IV), the estimated spatial-lag coefficient, $\hat{\rho}$ =0.3897, exceeds the true ρ=0.2, by 0.1897, or almost 100%, with an estimated standard error of just 0.0104. This *asymptotic bias* will not decrease as T grows large. By contrast, the $\hat{\beta}$ estimate is within one standard error of its true value. Conversely, if one includes fixed period-effects, and T is non-infinite, the $\hat{\rho}$ estimate incurs a *spatial Hurwicz bias* downward, paralleling the *LSDV* estimator discussed extensively in the econometric literature on dynamic-panel models (Hsiao 1986, Baltagi 1995). Analysis V (dummy coefficients omitted) shows the estimated $\hat{\rho}$ as 0.1156, or almost 50% lower than the true ρ, with an estimated standard error of 0.0262. This *spatial Hurwicz bias* is a small-sample property that decreases with T. Still, the $\hat{\beta}$ estimate is (barely) within a standard error of its true value; because **X** remains spatially orthogonal, any bias in $\hat{\beta}$ is small.

For analysts who have a substantive interest in spatial relationships, this apparent no-win situation is troubling. How can one obtain good estimates of both β and ρ? The dynamic panel models literature has proposed several alternatives to the *LSDV* estimator: a two-stage and *FIML* estimator (Anderson and Hsiao 1981), a corrected *LSDV* estimator (Kiviet 1995), and a *GMM* estimator (Arellano and Bond 1991), which last has begun to appear in some political-economy applications (**XXXX**). We expect similar approaches to bear fruit in the spatial context and will explore that conjecture in this project.

2. Spatially Correlated Disturbances with Spatially Correlated (Exogenous) Regressors

With spatially correlated disturbances but spatially orthogonal **X**, obtaining efficient and unbiased $\hat{\beta}$ estimates raises mostly surmountable challenges. When both regressors and disturbances correlate spatially, however, the difficulties increase appreciably. Formally,

$$6. \qquad \mathrm{var}(\varepsilon) = \sigma^2\mathbf{C} \;\; \textit{and} \;\; \mathrm{var}(\mathbf{X}) = \psi^2\mathbf{W}$$

Now, the variance-covariance matrix **var[X]** contains non-zero elements off the diagonal. The covariance in **X** across units arises from spatial diffusion. Again, we use the model $y_{n,t} = .2\mathbf{s}_t\mathbf{w}_n + .8x_{n,t} + \eta_t + \upsilon_{n,t}$ to generate a *30x20* dataset and conduct a one-shot simulation. If the analyst does not account for the spatial relationships in any way (Table 2a, Analysis I), the $\hat{\beta}$ estimate will be highly inefficient (and biased upward, we expect). Note that the estimated standard error, which is (also) biased, exceeds 0.15.

a. Estimating β with Simple Spatial Proxies (Table 2, Analyses II-III)

Can one improve the $\hat{\beta}$ estimate's efficiency (*i.e.*, reduce its standard error) simply and innocuously? Unfortunately: no, because either simple proxy, spatial indicators or spatial-lag averages, ultimately offers an imperfect substitute for the true diffusion and shock variables. Because **X** correlates spatially here, an omitted-variable (misspecification) problem arises from introducing the spatial proxies, which will almost always be endogenous, *i.e.*, correlated with the disturbance. If the *spatial-diffusion* matrix is symmetric and period effects omitted, the $\hat{\rho}$ estimate of *spatial diffusion* is biased upward because omitted *common shocks* correlate positively with the spatial lag. In the simulation, $\hat{\rho}$ exceeds ρ by 0.6481 (Analysis II). Moreover, $\hat{\beta}$ understates β by approximately 25% because the spatial lag "steals explanatory power" from **X** (Achen 2000 discusses a similar effect of temporal lags). Conversely, with spatial lags omitted and only time dummies included (Analysis III), the omitted lag correlates positively with the included **X,** and the $\hat{\beta}$ estimate is biased upward, by +0.0721 with an estimated standard error of 0.039 in the simulation.

Table 2a: Spatially Correlated Disturbances and Regressors (Simulated Data)

(True Model: $y_{n,t} = .2 * \mathbf{s_t w_n} + .8 * x_{n,t} + \eta_t + \upsilon_{n,t}$)

Variable	Analysis I	Analysis II	Analysis III	Analysis IV	Analysis V
X	0.9436**	0.59042**	0.87214**	0.4402**	0.7643**
	(.15084)	(.03595)	(.03861)	(.05234)	(.03404)
$\bar{y}_{\sim n,t}$		0.84807**			
		(.02545)			
SW				0.3149**	0.1741**
(Spatial Lag)				(.01560)	(.01202)
Fixed Period Effects	No	No	Yes	No	Yes
No. of Observations	600			600	600

**p-value < .01, *p-value < .05

Table 2b: Modeling Unemployment with Spatial and Temporal Lags (OECD, 1966-1990)

Variable	Analysis I	Analysis II	Analysis III
Unemployment$_{t-1}$	0.9179**	0.0309	0.0562*
(Temporal Lag)	(.01794)	(.02179)	(.02863)
Unemployment$_{\sim n,t}$		0.9697**	0.9448**
(Spatial Lag)		(.02256)	(.02952)
Fixed Unit Effects	Yes	Yes	Yes
Fixed Period Effects	No	No	Yes
No. of Observations	350	350	350
R^2	0.936	.990	.990

Data Source: Garrett (1998)

**p-value < .01, *p-value < .05

b. Getting Good Estimates of Both ρ and β (Table 2, Analyses IV and V)

The poor results in Analyses II-III do not derive from the spatial-diffusion matrix being unknown or from inability of time dummies and dependent-variable averages to enter these models simultaneously. Were **W** known, the analyst could include the <u>true</u> spatial lag in the regression model as in Analysis IV. As seen, however, the $\hat{\rho}$ coefficient remains overestimated, although the bias decreases radically from Analysis II. Moreover, $\hat{\beta}$ now underestimates β by about 50%. Again, this is the (incorrectly estimated) spatial lag robbing explanatory power from **X** (*a la* Achen 2000). Nor does unfolding the recursive process implied by the *spatial diffusion* –unit 1 affects units 2 through *N*, which in turn affect 1—produce more accurate *post-diffusion* $\hat{\beta}$ estimates. In Analyses II, IV, V, the corresponding *post-diffusion* $\hat{\beta}$ are .59/(1-.85)=3.88, and .44/(1-.31)=.64, and .76/(1-.17)=.92, only the last of which approaches the true .8/(1-.2)=1.

With **W** asymmetric, distinguishing *common shocks* from *spatial diffusion* becomes more feasible. *I.e.*, ability to discern *spatial diffusion* from *spatially correlated responses to omitted spatially correlated shocks* increases, we expect, with the difference in their incidence patterns and may depend, we conjecture, non-monotonically on the number of cross-sectional units. In our one-shot simulations, *e.g.*, an analyst could include both time dummies and the true spatial lag (or an appropriately parameterized model thereof). This spatial analogue to the *LSDV* estimator for dynamic panel-models reduces problems in estimating β to those arising in that parallel context. Namely, with both spatial lags and time dummies, $\hat{\rho}$ is biased downward in limited samples, but, without period indicators, $\hat{\rho}$ is biased upward asymptotically. In our simulation, the *spatial Hurwicz bias* is relatively small in absolute terms (Analysis III); $\hat{\rho}$ underestimates ρ by 0.0259, yet still by more than two estimated standard errors of 0.0120. Also, $\hat{\beta}$ underestimates β by about one standard deviation. Again, to get good estimates of both $\hat{\rho}$ and $\hat{\beta}$, analysts will need to apply spatial analogues to the more-sophisticated methods mentioned above.

As one final illustration of this *Hurwiczian* dilemma, consider some actual unemployment data for OECD countries between 1966 and 1990. Most analysts agree that unemployment rates persist strongly over time, and none would argue that the unemployment rate at time *t-1* has no effect on the rate at time *t*. Regressing annual unemployment rates on their temporal lag strongly supports this belief (Table 2b, Analysis I). The estimated temporal-lag coefficient is 0.9179, which is likely too large (although strong hysteresis may indeed plague unemployment). Now introduce an asymmetric spatial-lag to the model; specifically, include the predicted unemployment rate from an auxiliary regression of unemployment on the rates in other countries that year as a regressor in the model. With this spatial lag, the coefficient on the temporal lag almost vanishes, dropping to 0.0309, and becomes statistically insignificant (Analysis II). The omission of fixed period-effects in Analysis II causes some of this, biasing the spatial-lag estimate upward. Given the asymmetric spatial-lag, the model can accommodate common, fixed period-effects. With these included, the estimated temporal-lag and spatial-lag coefficients rise and decline, respectively, as expected; although the temporal-lag estimated remains small, it at least regains statistical significance. Notice from this example and those above that a spatial *diffusion process* also implies a spatial analogue to the *long-run multiplier* in dynamic models. In a model with a simple temporal lag, a coefficient of β on X and of ρ on y_{t-1} implies a long-run effect of a permanent 1-unit increase in X of $\beta/(1-\rho)$. The symmetric spatial-diffusion parameter on $\bar{y}_{\sim n,t}$ similarly implies that the post-diffusion total-impact of an exogenous unit-change in X, all of which accrues instantaneously given the contemporaneousness of the spatial lag, is $\beta/(1-\rho)$. Thus, in Analyses II-III, the *post-diffusion coefficients* on y_{t-1} are $\hat{\beta}_{y(t-1)}\big/(1-\hat{\rho}_{\bar{y}(\sim n,t)})$, or 1.0198 and 1.0181 respectively: now likely too large once again (in fact, explosive, but, again, such full hysteresis, *i.e.*, unit roots are possible). Our study will seek to offer some guidance in complicated, but common, contexts such as these with both spatial and temporal dynamics prominent.

3. Spatially Correlated and Endogenous Regressors

Finally, we consider the case of spatially correlated <u>and</u> endogenous regressors. Formally,

$$7. \qquad \mathrm{var}(\mathbf{X}) = \psi^2\mathbf{C} \quad \textit{and} \quad \mathrm{cov}(\mathbf{X},\upsilon) = \xi^2\mathbf{I}$$

The variance-covariance matrix **var[X]** is *NxN* with ψ^2 as the diagonal and non-zero elements off-diagonal. The variance-covariance matrix **cov[X,υ]** is *NxN* with ξ^2 as the diagonal and zeros off-diagonal. Here, we use a different model, $y_{n,t} = 0.8 \cdot x_{n,t} + \upsilon_{n,t}$; $x_{n,t} = \eta_t + 0.9\upsilon_{n,t}$, to generate the *30x20* dataset in which to conduct our one-shot simulation exercise. The model assumes that the disturbances to $y_{n,t}$ and $x_{n,t}$ correlate, reflecting the endoegneity of the regressors, and that the spatial correlation in **X** is caused by the *common shocks*. The project will add consideration of *spatial diffusion* among the **X** also.

Table 3: Spatially Correlated and Endogenous Regressors (Simulated Data)

	(True Model: $y_{n,t} = 0.8 * x_{n,t} + \upsilon_{n,t}$, $x_{n,t} = \eta_t + .9 * \upsilon_{n,t}$)	
Variable	OLS	2SLS (Instrument: $\overline{x}_{\sim n,t}$)
X	1.0833**	0.8168**
	(.01961)	(.03436)
Constant	0.0800*	.06658
	(.03610)	(.04125)
No. of Observations	600	600

**p-value < .01, *p-value < .05

a. Using Spatial Lags as Instruments: The Ideal Case

With endogenous variables among the right-hand-side terms in a regression equation, analysts might be able to leverage existing spatial interdependence to achieve better estimates of the effects of these variables. In particular, the endogenous variables from *other* cross-sectional units might provide at least near-valid instruments for the endogenous variables in each country's right-hand-side. As noted, political economists and other social scientists occasionally apply just such expedients, and, more frequently, simply assume the exogeneity of *external conditions* when using them directly as controls or substantive right-hand-side variables. With spatially orthogonal (independent) disturbances, and with the spatial correlation in **X** attributable to *common shocks* only, the average value of the independent variable in other spatial units each time period, $\bar{x}_{\sim n,t}$, correlates with the exogenous component of $x_{n,t}$ only and so works quite well as an instrument. Indeed, under these conditions, the *TSCS* structure of the data *automatically* offers a *perfectly exogenous* spatial instrument! Table 3 illustrates, comparing *OLS* (Analysis I) with *2SLS* using $\bar{x}_{\sim n,t}$ to instrument for ***X*** (Analysis II). Because $x_{n,t}$ correlates positively with the disturbances, the *OLS* $\hat{\beta}$ overestimates β (here, substantially). Conversely, the *2SLS* $\hat{\beta}$ estimate, which leverages $\bar{x}_{\sim n,t}$ for identification, is unbiased and within one-half of a standard error from the true β.

b. Using Spatial Lags as Instruments: The (Unfortunately) Usual Case in Practice

Unfortunately, we expect this simple and readily available instrument far more often to be imperfect, *i.e.*, to produce biased and inconsistent estimates because it too is endogenous; *i.e.*, it provides at best a *quasi-instrument* in Bartels' (1991) terms. The instrument in our example was perfect because the source of spatial correlation in **X**, which is what bought the instrument power in the first stage, was solely a spatially shared and *exogenous* shock, which did not therefore jeopardize the identification leverage of the instrument in the second stage. More commonly, however, **X** will correlate spatially *via* both *common shock* and *spatial diffusion*, just as does **Y**, especially in cases where one is already entertaining the notion that **X** and **Y** have the kind of conceptual symmetry implied by the definition of endogeneity: **X**↔**Y**. In fact, one may well suspect what is sadly usual regarding instruments: that the better $\bar{x}_{\sim n,t}$ predicts $x_{n,t}$, the more likely it is to be more strongly endogenous. Therefore, here as elsewhere, researchers must face *Bartels' dilemma*: a tradeoff between bias and efficiency whose terms depend on the ratio of the correlation of instruments with that for which they instrument, which is estimable, to the correlation of the true stochastic component with the instruments, which is not. In some cases the analyst may achieve smaller-*MSE* estimates by *OLS* than with weak or insufficiently exogenous *quasi-instruments* in *2SLS*. Recognizing that simple spatial-instruments such as the one suggested above will far more likely be imperfect *quasi-instruments* than perfect ones, our project seeks to characterize the pattern of spatial and other causal arrows, $\mathbf{X}\leftrightarrow\mathbf{Y}$, $\mathbf{X}_i\leftrightarrow\mathbf{X}_j$, and $\mathbf{Y}_i\leftrightarrow\mathbf{Y}_j$, that determine the terms of *Bartels' tradeoff* in this context. Again, analysts viewing spatial relationships as *substance* may prefer more-sophisticated techniques for modeling the potentially complex spatial interdependence and endogeneity in their data. We will explore several such methods, including some developed in the endogenous-choice/quasi-experimental literature (Heckman 1978, Achen 1986).

IV. Outline of the Methodological Project (and some Early Results)

As noted at the start, this project addresses both social-science researchers directly interested in spatial relationships (*spatial substance*) and those primarily concerned to make optimal inferences regarding other substantive relationships given spatial dependence (*spatial nuisance*). Our approach, broadly speaking, is to build from analogies to similar, better-explored issues arising from temporal dependence, and through analytic derivation and Monte Carlo experimentation. We will, as noted, tackle six tasks thusly:

(1) detail the conditions under which failing to model spatial dependence or relegating its role to standard-error adjustment biases other coefficient estimates or mostly induces *mere* inefficiency, exploring magnitudes of these biases and inefficiencies under differing degrees and natures of spatial dependence;

(2) distinguish *spatial diffusion* from spatially correlated responses to omitted spatially correlated

common shocks conceptually, and evaluate alternative approaches to making such distinctions empirically;

(3) compare the properties of simple proxies for fuller models of *spatial diffusion* or omitted spatially correlated *common factors*—*e.g.*, spatial dummies or symmetric spatial-lags comprised of averages of other cross-section units' dependent variables each time-period—to each other with and without *PCSE*, to *PCSE* alone, and to alternative techniques for estimating fuller or the complete model;

(4) explore the properties of leveraging spatial dependence to aid identification of endogenous systems by using spatial lags as *quasi-instruments* (Bartels 1991) under differing magnitudes of the relevant spatial correlations and causal arrows: $\mathbf{X}\leftrightarrow\mathbf{Y}$, $\mathbf{X_i}\leftrightarrow\mathbf{X_j}$, and $\mathbf{Y_i}\leftrightarrow\mathbf{Y_j}$ (or $\boldsymbol{\varepsilon}_\mathbf{i}\leftrightarrow\boldsymbol{\varepsilon}_\mathbf{j}$).

(5) develop and explore the properties of several parametric, semi-, and non-parametric methods to test for and gauge spatially correlated disturbances, in the presence or absence of spatial lags in the model;

(6) create and disseminate statistical-software algorithms in widely used software to implement the techniques explored and, where potentially useful, pedagogical modules to help teach them.

a. Tests for and Measures of Spatial Correlation

We have discussed the first four tasks above, and the last is self-explanatory. Regarding tests for and statistical measures of spatial correlation, notice that so far we have assumed that analysts working with *TSCS* data know they correlate spatially. Such data certainly will typically exhibit spatial correlation, but analysts will still want to gauge the magnitude of (*i.e.*, diagnose) this problem before treating it, especially if the bias-efficiency-complexity tradeoffs we mentioned manifest as frequently and prominently as we expect. For example, we expect that, just as including lagged dependent variables as regressors induces bias if residuals retain temporal correlation, including spatially lagged dependent variables as regressors induces bias if residuals retain spatial correlation. Accordingly, this project will develop several parametric, semi-parametric, and non-parametric tests for (gauges of) spatial correlation and provide analytic and experimental results documenting their large- and small-sample properties. We will include among these:

(1) Lagrange multiplier tests, derived by analogy to White's heteroscedasticity (*i.e.*, non-constant variance) test, which would regress $e_{i,t} \cdot e_{j,t}$ ($e \equiv$ estimated error) rather than White's $e^2_{i,t}$ on the corresponding products and, as feasible, cross products of the elements of $\mathbf{X_{i,t}}$ and $\mathbf{X_{j,t}}$ rather than White's products and, as feasible, cross products of the elements of $\mathbf{X_{i,t}}$. We have confidence in the analogy's soundness: as with heteroskedasticity, the sample pattern of which $e^2_{i,t}$ estimates, the spatial correlation, the sample pattern of which $e_{i,t} \cdot e_{j,t}$ estimates, that most imperils standard regression statistics will have a form whose pattern relates to the corresponding products and cross-products (*i.e.*, moments) of the independent-variable matrix. Moreover, we expect this test to retain validity with or without spatial lags in the model as does White's. However, the very large number of products and cross products of $\mathbf{X_{i,t}}$ and $\mathbf{X_{j,t}}$ ($1xK$ vectors) may severely limit feasibility (even ignoring cross-products, at some cost, to limit this concern), and the test statistic's small-sample/large-sample properties may/will prove difficult to simulate/derive.

(2) Lagrange multiplier (*LM*) tests derived by analogy to the standard *LM* test for serial correlation, which would regress $e_{i,t}$ on $e_{j,t}$, rather than $e_{i,t}$ on $e_{i,t-1}$, controlling for the other variables of the model. Like that inspired by White, this test should retain validity with or without spatial lags as its temporal analogue does. It should also prove more-widely feasible, and its large- and small-sample properties easier to derive and to simulate. However, such tests cannot distinguish the more perilous forms of spatial correlation where the spatial-dependence pattern correlates with the moments of the **X** matrix from the lesser forms.

(3) Breusch and Pagan's (1980) *LM* test, which is described in more detail below where we begin the Monte Carlo assessments of this test's size. However, although this test's asymptotic properties are known as its validity with and without spatial lags in the model, its small-sample properties are unknown, and it unfortunately cannot distinguish more from less perilous patterns of spatial correlation.

(4) Likelihood-ratio tests as described in Greene (1997: 661), which compare the sums of logged diagonal elements of the residual variance-covariance matrix under a restriction holding all off-diagonal elements to zero with the log of the unrestricted variance-covariance matrix's determinant, to test the null hypothesis that the unrestricted matrix's off-diagonals are all zero. Determinants of diagonal matrices are just the products of their diagonal elements, so, if restricted and unrestricted variance-covariance matrices are equal, the sum of the logged diagonal elements and the log of the determinant will be equal. The sum of the restricted matrix's logged diagonals will exceed the log of the unrestricted matrix's determinant by more the larger are the absolute values of the off-diagonal elements in the unrestricted matrix. Under the null hypothesis, T times this difference is distributed chi-squared with $N(N-1)/2$ degrees of freedom, which, notice, being equal to the number of sample covariances, depends only on N.

(5) Some semi- and non-parametric tests and measures based on groupings of residuals, including ones based on the familiar *Durbin-Watson* statistic, various spatial versions of autocorrelation and partial-autocorrelation correlograms, and *Ljeung-Box Q*. Some of these will not be robust to inclusion of spatial lags in the model, and their generally weaker structural assumptions may weaken them as tests for specific patterns of spatial dependence, such as that termed *most-perilous* above. However, conversely, we expect that their looser structure might broaden the range of patterns that might register with the test or measure.

Adding the development and evaluation of these tests and measures to the other five tasks, our research program will undertake six tasks or sets of questions in total, addressing each (a) by analogies to similar, more-familiar temporal-dependence issues and by analytic derivation and Monte Carlo experimentation, (b) for three alternative types of spatial interdependence—disturbances only; disturbances and exogenous regressors; and disturbances endogenous regressors—and (c) for two audiences—those for whom spatial dependence is solely *nuisance* potentially jeopardizing estimates of other substantive quantities of primary interest and those for whom the spatial dependence itself holds more-central *substantive* interest.

b. Tests for and Measures of Spatial Correlation: The Size of Breusch and Pagan's LM Test

The size of a statistical test is its probability of a *Type I Error*, *i.e.*, of falsely rejecting a null hypothesis that is, in fact, true. In this section, we conduct Monte Carlo experiments to examine the size of Breusch and Pagan's (1980) *LM* test in small samples. Each experiment is defined by its sample dimensions, *i.e.*, the number of units and periods. The results presented derive from 1000 independent trials. We use the following model to create a sample for *N* units over *T* periods:

$$8. \qquad y_{i,t} = x_{i,t} + \varepsilon_{i,t}$$

Variables are indexed $i = 1...N$ and $t = 1...T$ to identify each of the sample's *NT* observations. The exogenous variables x and ε are independent draws from a standard normal distribution. Hence, the data exhibit no spatial correlation. After generating the sample, we estimate the following model by *OLS*:

$$9. \qquad y_{i,t} = \beta x_{i,t} + \gamma \bar{y}_{\sim i,t} + \varepsilon_{i,t}$$

The variable $\bar{y}_{\sim i,t}$ is the average *y*-value for unit *i*'s sample counterparts, $\{j\}$, at time *t*, this being one way some (*e.g.*, Franzese 2002) have attempted to address (some of) the spatial correlation in their *TSCS* data. Breusch and Pagan (1980: 247) calculate their *LM* statistic as:

$$10. \qquad LM = T\sum_{i=2}^{N}\sum_{j=1}^{i-1} r_{ij}^2$$

with r_{ij} the ij^{th} residual correlation. Under the null, *LM* is distributed asymptotically chi-squared with *(N·(N-1))/2* degrees of freedom. A $\chi^2_{N\cdot(N-1)}$ distribution thus determines the critical values for our test.

For now, our primary interest is how this *LM* statistic performs under different sample dimensions (*NxT*). We consider several *NxT* dimensions common in political economy research: *T*'s of 20, 30, or 40 (years) by *N*'s of 5, 10, 15, 20, or 30 (countries). We conduct three experiments for each *T* using different *N*'s, giving nine in total: *5x20, 10x20, 15x20, 5x30, 10x30, 20x30, 10x40, 20x40,* and *30x40*.

Table 4 reports the results for each experiment. We focus on the *95th* percentile of the relevant chi-squared distribution as the critical value for comparison. Not surprisingly, the experimental size of the *LM* test for spatial correlation in small samples always exceeds the asymptotic size of the test: *0.05*. In some cases, it more than doubles this 5% size. The *LM* test may help diagnose spatial correlation, but, as always, analysts must use care in interpreting borderline results suggesting correlation, especially in small samples where it rejects appreciably more often than warranted. One result of the experiments is surprising: for a given *T*, smaller *N* does not always produce more-accurate test-sizes (*e.g.*, test size seemed truer at *10x30* than at *5x30*); nor does increasing *T* for a given *N* always yield truer test-size (*e.g.*, test size is truer at *10x30* than at *10x40*). This suggests an optimal *NxT* ratio for test accuracy may exist, but, generally, this *LM* statistic performs reasonably well, though *p*-values based may understate true probabilities of *Type I Errors* in small samples. As we shall see below, however, the test easily reveals strong spatial correlation where it is present, although whether this power is accurate (*i.e.*, avoidance of *Type II Errors*) at smaller degrees of spatial correlation and/or in smaller samples remains untested (for now).

Table 4: Size of Breusch and Pagan (1980) LM Test under Null Hypothesis of No Spatial Correlation

N:	***5***	***10***	***15***	***5***	***10***	***20***	***10***	***20***	***30***
T:	***20***	***20***	***20***	***30***	***30***	***30***	***40***	***40***	***40***
Trials:	*1000*	*1000*	*1000*	*1000*	*1000*	*1000*	*1000*	*1000*	*1000*
Percentiles									
1%	3.68	28.29	82.22	3.12	29.03	158.60	28.79	154.57	381.87
5%	5.04	34.59	91.22	4.70	32.88	168.14	33.04	164.82	399.21
10%	6.20	36.87	94.10	5.87	35.84	174.75	35.59	172.09	409.12
25%	8.19	41.20	100.90	8.27	40.64	184.67	40.86	182.10	424.64
50%	11.23	47.10	110.83	11.04	46.51	196.48	47.26	195.81	446.34
75%	14.60	53.15	121.48	14.19	53.30	211.30	54.23	209.05	466.71
90%	18.20	59.50	130.99	18.23	60.12	224.42	60.64	224.03	487.62
95%	**20.62**	**65.06**	**138.49**	**20.58**	**64.04**	**230.62**	**65.21**	**233.83**	**497.55**
99%	27.96	76.46	150.91	27.62	74.87	248.61	74.79	251.69	521.02
Degrees Freedom:	10	45	105	10	45	190	45	190	435
Chi-Squared (95%):	***18.31***	***61.66***	***129.92***	***18.31***	***61.66***	***223.16***	***61.66***	***223.16***	***484.63***
Size of LM Test:	***0.095***	***0.08***	***0.113***	***0.096***	***0.074***	***0.108***	***0.086***	***0.104***	***0.116***

V. Modeling Spatial Dependence: An International Tax-Competition Example of Asymmetrical Diffusion

1. Globalization, Tax Competition, and Converge: Literature Overview

In theory, strong inter-jurisdictional competition undermines the tax-policy autonomy of individual tax authorities, inducing tax rates to converge, especially those levied upon more-mobile assets. Such inter-jurisdiction competition intensifies as capital becomes more liquid and more mobile across borders. Indeed, many scholars of domestic and international fiscal-competition (*e.g.*, Zodrow and Mieszkowski 1986, Wilson 1986, Wildasin 1989; Oates 2001, Wilson 1999 review) expect such intense inter-jurisdiction competition to engender a virtually unmitigated race to some (ill-defined: see below) bottom. As a central exemplar, most scholarly and casual observers see the striking post-1970s rise in international capital mobility and steady postwar increase in trade integration as forcing welfare/tax-state retrenchment and a shift in tax-burden incidence from relatively mobile (*e.g.*, capital, especially financial) toward more immobile (*e.g.*, labor, especially less-flexibly-specialized types[3]). Growing capital-market integration and asset mobility across jurisdictions enhances such pressures, the argument holds, by sharpening capital's threat against domestic governments to flee "excessive and inefficient" welfare/tax-systems.

Several notable recent studies of the comparative and international political economy of policy change over this period challenge these claims however. First, empirically, that *globalization* in general and capital mobility in particular has actually constrained public policies in general and capital-tax policy is contested. Hines (1999), after reviewing the empirical economics literature, concludes that national tax-systems affect the investment location decisions of multinational corporations and firms do seize opportunities for tax avoidance. Rodrik (1997), Dehejia and Genschel (1999), Genschel (2001), and others argue that this has increasingly constrained governments' policy-latitude in recent years. Quinn (1997), Swank (1998, 2002), Swank and Steinmo (2002), Garrett and Mitchell (2001), and others, however, do not find these trends to have constrained governments' tax policies much or at all. The theoretical explanation for such results, occasionally implicit, seems that other cross-national differences (*e.g.*, commercial, regulatory, and other policy, labor-market institutions, intermediate-supply availability, final-market proximity, *etc.*: Hines 1999: 308) also importantly affect investment-location decisions, affording governments some room to maneuver. Moreover, other factors than capital mobility affect governments' tax policies. Swank (2002: esp. p. 252-6, Table 7.1) argues, *e.g.*, that capital and corporate tax rates are a function of funding requirements of programmatic outlays, macroeconomic factors like inflation and economic growth, and partisan politics. Controlling for such factors, he finds little relationship between capital mobility and taxation.

[3] Unskilled labor is usually relatively mobile within (national) jurisdictions but highly immobile across jurisdictions, especially those borders delineating strongly differentiated ethnic, linguistic, religious, and cultural societies. Some types of skilled labor is highly specialized into specific productive activities, which may limit intra- and inter-jurisdictional mobility; other types, some *human capitalists*, *e.g.*, may be relatively mobile across jurisdictions.

On closer inspection, these recent challenges to *globalization-induces-welfare/tax-state-retrenchment* arguments have four distinct bases. Garrett (1998) argues that certain combinations of left government with social-welfare, active-labor-market, coordinated-bargaining, and related policies can be as or more efficient than neoliberal state-minimalism and conservative government and, therefore, that capital will not flee such efficient combinations. Boix (1998) argues that public human- and physical-capital-investment strategies comprise an alternative to neoliberal minimalism that is sufficiently efficient macroeconomically to attract and retain capital and politically effective enough to maintain left electoral-competitiveness. Hall, Soskice, and colleagues (2001) argue that complex national networks of political-economic institutions confer *comparative* advantages in differing productive activities, which, as Mosher and Franzese (2002) elaborate, implies capital mobility and trade integration could (*if international tax-competition remains sufficiently muted*: see below) spur institutional and policy specialization—here, cross-national variation in welfare/tax systems—rather than convergence or global retrenchment. These views fundamentally question whether international economic integration actually creates economic pressures to retreat from welfare/tax-state commitments (or at least whether all aspects of globalization do so, so strongly: see below).

Swank's (2002) argument, that the institutional structures of the polity and of the welfare system itself shape the domestic policy-response to integration, represents a fourth basis for challenge. His view does not fundamentally challenge claims of the exclusive or superior macroeconomic efficiency of neoliberal minimalism but, rather, stresses the primacy of domestic political conditions—the policymaking access, cohesion and organization, and relative power of contending pro- and anti-welfare/tax interests—in determining the direction and magnitude of welfare/tax-policy reactions to global economic integration. Specifically, he argues and finds—in *TSCS* statistical analysis supplemented by thorough qualitative case explorations of generous welfare states and briefer explorations of less-generous systems (respectively, the *universal* and *conservative*, and the *liberal*, systems of Esping-Andersen (1990))—that inclusive electoral institutions, social-corporatist interest-representation and policymaking, centralized political authority, and universal welfare systems relatively favor the political access and capacity of pro-welfare/public-policy interests and bolster supportive social norms in the domestic political struggle over the policy response to integration. The opposite conditions favor anti-welfare/tax interests and norms in this struggle. Capital mobility and globalization therefore induce increased welfare/tax-state largesse in previously generous states and retrenchment in tight ones: *i.e.*, divergence not convergence. Swank's approach is, thus, the most directly and thoroughly political of these critiques. It is also perhaps the most thoroughly explored empirically, offering comparative-historical statistical analyses against six alternative versions of the globalization-induces-retrenchment thesis: a simple version (a regression including one of five capital-openness measures), and five others he terms the *run-to-the-bottom* (capital openness times lagged welfare-policy), *convergence* (capital openness times the gap from own to cross-country mean welfare-policy), *nonlinear* (capital openness and its square), *trade-and-capital-openness* (their product), *capital-openness-times-fiscal-stress* (deficits times capital openness), and *capital-flight* (net foreign direct investment) versions. He finds little support for any globalization-induces-retrenchment argument, and, indeed, some indications that capital mobility tends on average to enhance welfare effort (perhaps supporting those stressing its effect in increasing popular demand for social insurance against global risks).[4]

Such critiques underscore that the *bottom* toward which globalization generally and capital mobility specifically may push tax-competing states may not be that of neoliberal minimalism.

[4] Franzese (2003) offers a more-complete review, including some more-critical suggestions for further research.

Insofar as alternative economic advantages allow some states to retain higher tax rates or restraining political conditions prevent some from reaching that neoliberal minimum, the competitive pressures on all states diminish, more so, of course, the more economically integrated and important are those states that retain such maneuvering room or suffer such constraints due to their own domestic political-economic conditions. Moreover, if, as Mosher and Franzese (2002) suggest, national economic-policy differences contribute to *comparative* advantages—which, if they do, they do regardless of their *absolute* efficiency—then both trade and global *fixed-capital* integration would enhance economic pressures toward specialization, *i.e.*, divergence not convergence. From this view, international *liquid-capital* mobility alone, through the tax-competition it engenders, produces whatever race to the bottom may occur. In this case, interestingly, such competitive races would occur regardless and independent of the efficiency of the tax systems in question or of the public policies they support. Furthermore, zero sets no *bottom* to such tax-cut races. On this competition for liquid capital alone, governments would always have incentives to cut taxes further, perhaps deep into subsidy; only their abilities to tax other less liquid and/or mobile assets and to borrow limit the race.

Notice, however, that international tax-competition arguments, in any of their conventional forms and throughout each of the many critiques, imply spatial interdependence in the cross-national rates of capital taxation. Whatever pressures upon domestic policymaking may derive from rising (liquid) capital mobility, their nature and magnitude will depend on the constellation of tax (and broader economic) systems with which the domestic economy competes. We offer preliminary theoretical and empirical models of capital-tax competition that reflects this international interdependence next.

2. International Interdependence in Theoretical and Empirical Models of Capital-Tax Competition

a. A Stylized Theoretical Model of Capital-Tax Competition

To show how tax competition implies spatial interdependence, we draw on Persson and Tabellini's (2000, Chapter 12) theoretical framework. In brief, the model's essential elements are as follows. In two jurisdictions or countries, denote the domestic and foreign tax rates τ_K and τ_K^*. Individuals differ in their relative labor to capital endowment, denoted e^i. Capital can be invested in either jurisdiction, but foreign investment incurs a "mobility cost." Following Besley and Coate (1997), an elected citizen-candidate, a policymaker with endowment e^P, sets the tax rate to maximize his or her own welfare function. Running for office is costly and citizens choose whether to enter the election by an expected-utility calculation. The stages of the model are: 1) elections occur in both countries, 2) the elected citizen-candidates set their respective countries' tax rates, and 3) all private economic decisions are made.

The best-response functions are $\tau_K = T(e^P, \tau_K^*)$ and $\tau_K^* = T^*(e^{P*}, \tau_K)$. In words, the domestic (foreign) capital-tax rate depends on the domestic (foreign) policymaker's labor-capital endowment and the foreign (domestic) capital tax rate—that is, capital taxes are spatially interdependent. The slope of these functions, $\frac{dT}{d\tau_K^*}$ and $\frac{dT^*}{d\tau_K}$, can be either positive or negative. An increase in foreign

Figure 1. Best Response Functions (Persson and Tabellini 2000, 334)

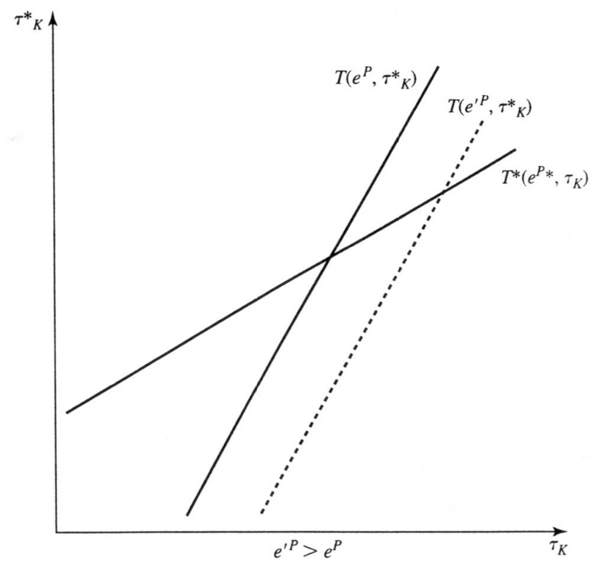


Figure 12.4

tax-rates induces a flow of capital into the domestic economy. The domestic policymaker may use the increased tax-base to lower tax-rates or to raise them to seize the greater revenue opportunities created by the decreased elasticity of this base. Figure 1 graphs these reaction functions under the assumption that $\frac{dT}{d\tau_K^*} > 0$ and $\frac{dT^*}{d\tau_K} > 0$. The illustrated comparative static shows an increase in the domestic policymaker's labor-capital endowment. This change shifts the function T outward, raising the equilibrium tax rate in both countries.

Even though tax-competition models, like Persson and Tabellini's, demonstrate that capital taxes are spatially interdependent, very few scholars have empirically modeled this interdependence directly.

b. International Interdependence and an Empirical Model of Capital-Tax Competition

We now propose and estimate a spatial error-correction model of tax competition reflecting the spatial interdependence that the theoretical model implies. First, though, we conduct the Breusch-Pagan *LM* test to gauge the statistical evidence of spatial correlation in our data: the Mendoza *et al.* (1997) capital tax-rates, as updated by Volkerink and de Haan (2001). We chose our sample of 12 OECD countries—Australia, Belgium, Canada, Finland, France, Germany, Italy, Japan, Sweden, Switzerland, UK, and US—from 1966-1996 to create the largest *balanced-panel* dataset possible.[5] The *LM* statistic for these raw capital-tax data is 734.33; with a critical value of 85.96, this clearly indicates strong spatial correlation and strongly suggests a spatial model. We estimate a spatial model that reflects the theory above in two steps.

First, we estimate the spatial interdependence of capital tax-rates by regressing each country's rate on each of its (sample) economic partners' tax-rates, a deterministic time trend, and the latter interacted with each of the former. These models represent a reduced form of the Persson and Tabellini model, allowing asymmetrical influence across countries. Jointly, they also represent an estimate of the *spatial-diffusion* matrix. We propose considering the predicted tax-rate in this regression as a sort of domestic *equilibrium* tax-rate that, given the set of tax-rates among its competitors, yields no further pressures from global tax-competition: a *competition-neutral* tax-rate. The interaction terms allow a country's *competition-neutral* tax-rate to change (linear-deterministically) over time as international capital mobility rises (Persson and Tabellini 1992). We then include the time *t-1* residuals from this model among regressors in the second-stage model

[5] As of now, all of the code we have written assumes balanced panels.

predicting the time *t* change in domestic capital tax-rates. This is the *spatial-error-correction*. Given our conception of the estimates at that stage as the *competition-neutral* rates, the residuals from stage one indicate the degree to which domestic taxes are *competition-nonneutral* or in *disequilibrium*. In the logic of tax-competition models like that above, the domestic policymaker should, conditional on the other political-economic factors in the stage-two model, adjust accordingly to regain neutrality. The coefficient on that error-correction term in the second stage will indicate how rapidly such adjustment occurs.[6]

The first-stage coefficient-estimates are also substantively interesting. Finding a country's capital tax-rate spatially independent or exogenous would be strong evidence against competition and for national policy-autonomy. To test the independence (autonomy) hypothesis, we implement the spatial equivalent of a Granger causality test (Freeman 1983), which is simply a joint *F*-test of the hypothesis that all the spatial coefficients are zero.[7] The *LM* statistic, reported above, strongly rejected the hypothesis of no spatial correlation in the *dataset*; these *F*-tests, which are country specific, yield equally unambiguous results. No country in our sample has spatially exogenous capital tax-rates. The only cases that fail to reject at the 0.01 level are the US (p=0.0388) and UK (p=0.0173), which seems substantively intuitive.[8]

Do the sorts of equilibrium relations identified in tax-competition models generate the spatial correlation of capital tax-rates across countries seen in stage one? To answer this question, we regress the time *t* change in the domestic capital tax-rates on the time *t-1* residuals from our first

[6] Notice that, whereas we have allowed each country to affect the competition-neutral tax-rate of others differently in stage one, we have, for now anyway, constrained all domestic policymakers to respond at an equal rate to net competition-nonneutrality.
[7] This suggests another of the tests for spatial correlation whose properties we intend to explore in our methodological project.
[8] Sweden is next-least significant, which also seems intuitive considering its relative leadership in Scandinavia and the relative concentration of inter-regional trade: the same considerations *vis-à-vis* the *Commonwealth* that, to greater degree, rendered the UK result intuitive. The US result intuitiveness rests on its overall economic leadership and greater trade outside our sample.

stage models. The idea, as noted, is that these residuals measure the degree of *disequilibrium* or *competition nonneutrality* in a country's tax rate. A large positive residual at time *t-1* means the country's tax rate is above its *competition-neutral* level, so we would expect a negative change in time *t* to restore *neutrality*. Conversely, a large negative residual at time *t-1* means the tax rate is below *neutrality* and we would expect a positive time-*t* change. Figure 2 presents these second-stage results. As expected, a negative and statistically significant relationship exists between the time *t-1* residuals in stage one and the time *t* change in the capital tax-rate.

Figure 2. Mendoza et al. Capital Tax Rate

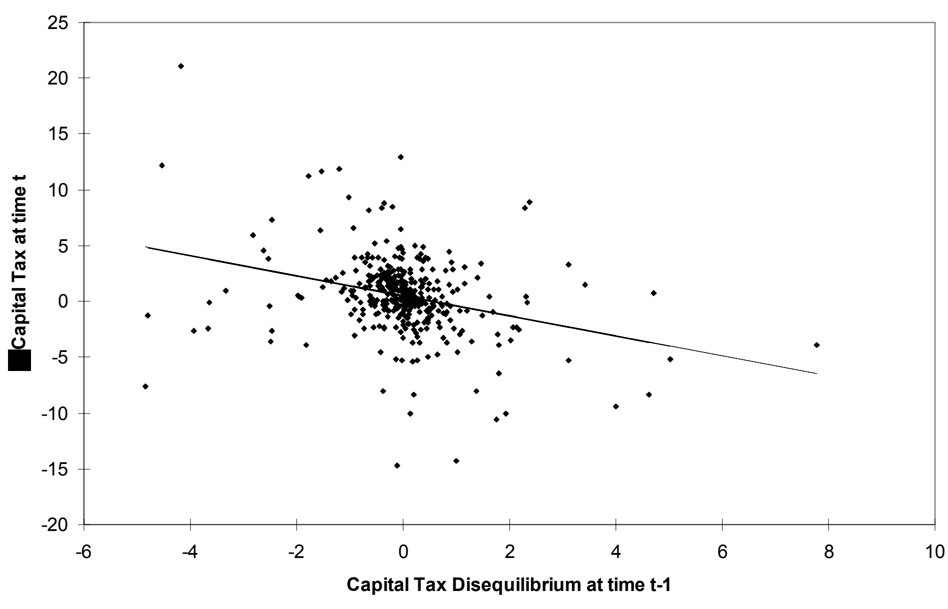
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.304558298
R Square	0.092755757
Adjusted R Square	0.090303745
Standard Error	3.369857385
Observations	372

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	429.5773936	429.5773936	37.82843509	2.00494E-09
Residual	370	4201.697354	11.35593879		
Total	371	4631.274747			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.499109924	0.174720086	2.856625909	0.004523457	0.155541089	0.84267876	0.155541089	0.84267876
CT Disequilibrium	-0.895075762	0.145529357	-6.150482509	2.00494E-09	-1.181244071	-0.608907453	-1.181244071	-0.608907453



Next, we offer some preliminary robustness checks. Some scholars criticize Mendoza *et al.* tax ratios, so we consider also the *effective average corporate tax-rates* of Devereux, Griffith, and Klemm (2002).[9] Their approach is similar to the well-known *cost-of-capital* method of King and Fullerton (1984).[10] Our results with Devereux *et al.* corporate tax-rates are very similar to those with Mendoza *et al*. No sample countries have tax rates independent of their (sample) economic partners' tax-rates, and the time *t-1* residuals from the stage one spatial-model predict time *t* changes in corporate tax-rates (Fig. 3).

Technically, the slopes of the best-response functions in Figure 1 are unidentified in our reduced form empirical setup. This is precisely the same problem one encounters when trying to estimate the slope of supply and demand curves with equilibrium price and quantity data only: an infinite number of curves could produce the same equilibrium outcome. This does not hinder our particular use of the estimates since identification is not required for prediction. In other words, our regressions offer good predictions of the *competition-neutral* capital tax-rate, which are all the error-correction model needs. Nevertheless, since the slopes of the best-response functions are of substantive interest, we also estimate our stage one spatial regression using two-stage least squares. That exogenous domestic factors partly determine each country's capital tax-rate seems likely (see, *e.g.*, Swank 2002). If so, these variables could supply instruments. We estimate our spatial regression treating each right-hand-side capital tax-rate as endogenous and using debt levels and government

[9] Mendoza et al. calculate their rates using the total tax payment as a proportion of some conventional measure of the tax base. For the capital tax rate, they use the operating surplus of the economy as the tax base. Devereux et al. point out that, if this measure were identical to the true tax base, as defined by the tax system, the Mendoza et al. rate would equal the statutory tax rate. Differences between the measured and true tax base reflect the fact that legislators deliberately define the tax base to be smaller or larger than the conventional base. Devereux et al. point out that current tax liabilities, particularly for corporations, reflect: 1) the history of investment, which determines allowances in the current period 2) tax liabilities in multiple jurisdictions, 3) the history of losses, which can be carried forward, and 4) the history of the tax system. In these ways, the Mendoza et al. rates are "backward looking" and unlikely to affect future investments decisions (Devereux et al. 2002, 468-9).

[10] These approaches infer the "cost of capital" from a net present value calculation, and then use this cost to compute effective tax rates (for details, see Devereux et al. 2002, 461).

partisanship (center of gravity) as exogenous instruments. Overall, the results for the spatial error-correction model do not change much (see Figure 4).

Figure 3. Devereux et al. Effective Average Corporate Tax Rate

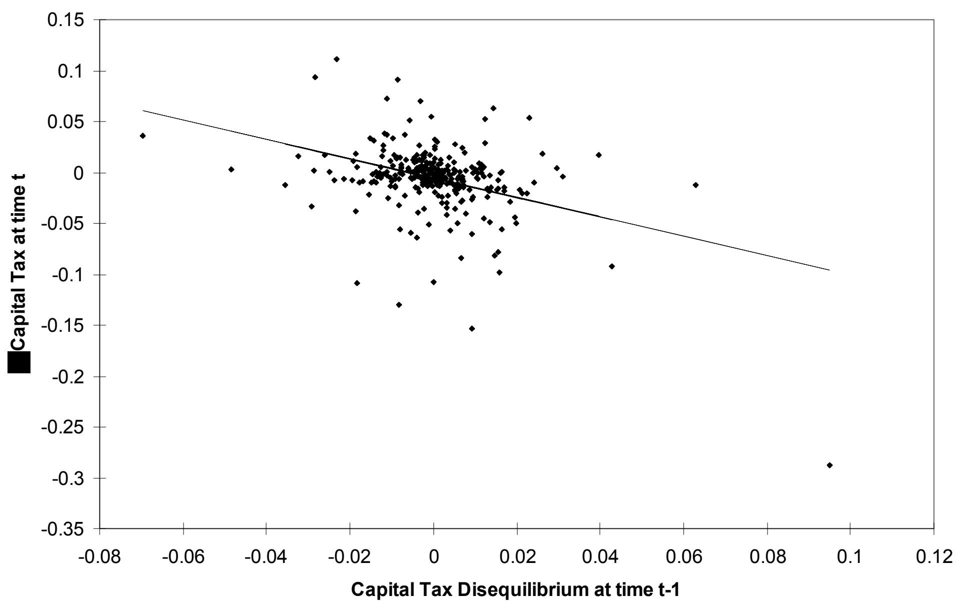
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.391956956
R Square	0.153630255
Adjusted R Square	0.150827707
Standard Error	0.0301021
Observations	304

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.049672631	0.049672631	54.81804785	1.32548E-12
Residual	302	0.273653208	0.000906136		
Total	303	0.323325839			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.005644645	0.001726558	-3.269305787	0.001202668	-0.009042254	-0.002247037	-0.009042254	-0.002247037
CT Disequilibrium	-0.950574682	0.128388008	-7.403921113	1.32548E-12	-1.203223145	-0.697926219	-1.203223145	-0.697926219



Finally, we re-estimate our original spatial-error correction model (see Figure 2) controlling for a number of additional variables and using standard methods for analyzing *TSCS* data. Specifically, we lag the dependent variable, include fixed effects, report *PCSE*, and control for lagged debt,

Figure 4. Mendoza et al. Capital Tax Rates (2SLS Results)

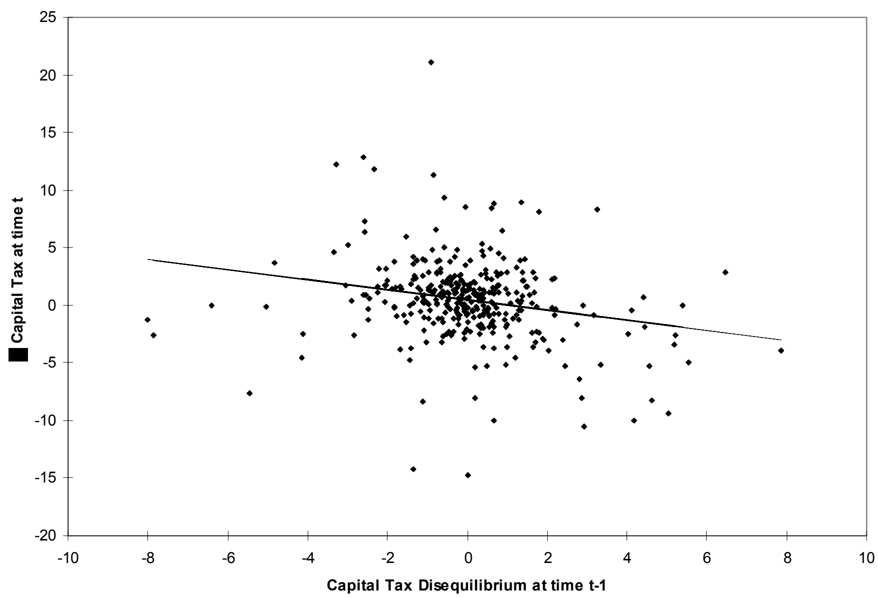
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.222588406
R Square	0.049545599
Adjusted R Square	0.046798621
Standard Error	3.502860754
Observations	348

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	221.3072494	221.3072494	18.03640145	2.78861E-05
Residual	346	4245.431578	12.27003346		
Total	347	4466.738828			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.471034484	0.187773046	2.508530879	0.012580273	0.101713828	0.840355141	0.101713828	0.840355141
Error	-0.441204132	0.103887818	-4.246928473	2.78861E-05	-0.645535466	-0.236872797	-0.645535466	-0.236872797



unemployment, real-GDP growth, trade openness, inflation, and government partisanship (factors Swank 2002 identified as important). Table 5 reports results. The lagged dependent variable, spatial error-correction term, and growth variable all receive correctly signed and statistically significant coefficients. The debt variable is correctly signed and borderline significant. Partisanship is also correctly signed, but insignificant.

Table 5. Multivariate Regression Results

	Coef.	PCSE	z	P>|z|
Lagged DV	.3169866	.0837531	3.78	0.000
Error Correction Term	-1.147844	.1958894	-5.86	0.000
Debt	2.200018	1.342113	1.64	0.101
Unemployment	-.1507454	.1171767	-1.29	0.198
GDP Growth	.4563691	.0802553	5.69	0.000
Trade Openness	3.835564	3.123282	1.23	0.219
Inflation	.0369919	.0724842	0.51	0.610
Partisan Center of Gravity	-.1430701	.160376	-0.89	0.372

VI. Conclusions

Most social scientists realize that the *TSCS* data they analyze are spatially interdependent—*i.e.*, cross-sectional variables correlate contemporaneously. Few, however, address the spatial relationships in their data seriously enough to model them further than employing *PCSE*. In this paper, we have argued direct modeling of spatial dependence is superior, even if one has little substantive interest in spatial relations. Direct modeling of spatial dependence increases efficiency and, in some cases, is necessary to avoid sizable bias and inconsistency in estimated non-spatial regressor coefficients. For comparative and international political economists, globalization makes these methodological issues increasingly central. Put simply, globalization causes spatial interdependence. Therefore, tools for diagnosing and modeling the types of spatial relationships found in *TSCS* data, particularly in political-economy datasets, need to be developed.

In this paper, we have suggested several possible tests for spatial correlation, and begun to analyze the small-sample properties of one: the Breusch and Pagan *LM* test. We find this test has good size properties in small samples, although p-values based on the χ^2 distribution may understate true probabilities of *Type I Errors*. The *LM* test has no difficulty revealing strong spatial correlation where it is present, however. Also, we have begun to explore a new way to model spatial *equilibrium* relationships—a *spatial error-correction* model—and applied this technique in an analysis of capital tax rates. That others have not modeled the spatial relationships in capital tax-rates is surprising given that most, if not all, theoretical models of tax competition imply spatial interdependence. Our analysis finds strong evidence of spatial correlation in OECD capital tax-rates. Moreover, this dependence seems to suggest the existence of *competition-neutral* capital tax-rates generating a sort of *equilibrium* cross-national relationship in rates of capital taxation that is driven, in part, by international capital mobility and tax competition.

References

Christopher H. Achen, *The Statistical Analysis of Quasi-experiments* (Berkeley, California: University of California Press, 1986).

Christopher H. Achen, 2000. “Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables.” Paper presented at the Annual Meeting of the Political Methodology Section of the American Political Science Association, UCLA, July 20-22.

Michael Alvarez, Geoffrey Garrett, and Peter Lange. 1991. “Government Partisanship, Labor Organization and Macroeconomic Performance.” American Political Science Review 85:539-56.

Theodore W. Anderson, Cheng Hsiao. 1981. “Estimation of dynamic models with error components.” Journal of the American Statistical Association 76:598-606.

Luc Anselin, *Spatial Econometrics: Methods and Models* (Boston: Kluwer Academic, 1988).

Manuel Arellano and Stephen Bond. 1991. “Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations.” Review of Economic Studies 58:277-297.

Badi H. Baltagi, *Econometric Analysis of Panel Data* (New York: John Wiley and Sons, 1995)

Lawrence Bartels, 1991. “Instrumental and ‘Quasi-Instrumental’ Variables.” American Journal of Political Science 35(3): 777-800.

Nathaniel Beck and Jonathan Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time Series Cross Section Models." Political Analysis 6:1-34.

Nathaniel Beck and Jonathan Katz. 1995. "What to do (and Not to Do) with Time Series Cross Section Data in Comparative Politics." American Political Science Review 89:634-47.

Nathaniel Beck, Jonathan Katz, and R. Michael Alvarez, Geoffrey Garrett, Peter Lange. 1993. "Government Partisanship, Labor Organization, and Macroeconomic Performance: A Corrigendum." American Political Science Review 87:945-8.

Carles Boix, *Political Parties, Growth, and Equality* (Cambridge University Press 1998).

Vivek H. Dehejia and Philipp Genschel, "Tax Competition in the European Union," Politics and Society 27 (1999)

Michael Devereux, Rachel Griffith and Alexander Klemm, 2002. "Corporate Income Tax: Reforms and Tax Competition," *Economic Policy* (October): 451-495.

Gosta Esping-Andersen. 1990. *The Three Worlds of Welfare Capitalism*, Princeton: Princeton University Press.

Robert J. Franzese, Jr. 1999. "Partially Independent Central Banks, Politically Responsive Governments, and Inflation," *American Journal of Political Science* 43(3): 681-706.

Robert J. Franzese, Jr. 2002. *Macroeconomic Policies of Developed Democracies*. Cambridge: Cambridge University Press.

Robert J. Franzese, Jr. 2003a. "Book Review: Duane Swank, *Global Capital, Political Institutions, and Policy Change in Developed Welfare States*. (New York: Cambridge University Press, 2002," *Policy Studies Quarterly*.

Robert J. Franzese, Jr. 2003b. "Multiple Hands on the Wheel: Empirically Modeling Partial Delegation and Shared Control of Monetary Policy in the Open and Institutionalized Economy," in review.

Geoffrey Garrett and Deborah Mitchell, "Globalization, Government Spending and Taxation in the OECD," European Journal of Political Research 39 (2001).

Geoffrey Garrett, *Partisan Politics in the Global Economy* (Cambridge University Press 1998).

Philipp Genschel, "Globalization, Tax Competition, and the Fiscal Viability of the Welfare State," Working Paper 01/01 (Cologne, Germany: Max Planck Institute for the Studies of Societies, 2001).

William Greene, Econometric Analysis. Third ed. (Upper Saddle River, N.J.: Prentice Hall, 1997).

Peter A. Hall, David Soskice, eds., *Varieties of Capitalism* (Oxford University Press 2001).

James Heckman, 1978. "Dummy Endogenous Variables in a Simultaneous Equation System" Econometrica 46(4): 931-959.

James R. Hines, "Lessons from Behavioral Responses to International Taxation," National Tax

Journal 52 (June, 1999).

Cheng Hsiao, *Analysis of Panel Data*. (New York: Cambridge University of Press, 1986)

Leonid Hurwicz, 1950. “Least Squares Bias in Time Series.” in Statistical Inference in Dynamic Economic Models. ed. T.C. Koopmans, pps. 365-383. New York: Wiley.

Ruth A. Judson and Ann L. Owen. 1999. “Estimating Dynamic Panel Models: A Guide for Macroeconomists.” Economic Letters 65: 9-15.

Jan F. Kiviet, 1995. “On bias, inconsistency, and efficiency of various estimators in dynamic panel data models.” Journal of Econometrics 68:53-78.

Enrique Mendoza, Assaf Razin, and Linda Tesar, “On the Ineffectiveness of Tax Policy in Altering Long-Run Growth: Harberger’s Superneutrality Conjecture,” Journal of Public Economics 66 (1997).

Enrique Mendoza, Assaf Razin, and Linda Tesar, “Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption,” Journal of Monetary Economics 34 (1994).

James M. Mosher and Robert J. Franzese, Jr., 2002, “Comparative Institutional and Policy Advantage,” *European Union Politics* 3(2):177-203.

Wallace E. Oates, “Fiscal Competition or Harmonization? Some Reflections,” *National Tax Journal* 54 (September 2001).

Torsten Persson and Guido Tabellini, *Political Economics: Explaining Economic Policy* (Cambridge, MA: The MIT Press, 2000).

Torsten Persson and Guido Tabellini, “The Politics of 1992: Fiscal Policy and European Integration” *The Review of Economic Studies* 59(4):689-701.

Dennis Quinn, “The Correlates of Change in International Financial Regulation” American Political Science Review 91 (1997)

Dani Rodrik, *Has Globalization Gone Too Far?* (Washington, DC: Institute for International Economics, 1997).

Duane Swank, “Funding the Welfare State: Globalization and the Taxation of Business in Advanced Market Economies,” Political Studies 46 (1998)

Duane Swank, *Global Capital, Political Institutions, and Policy Change in Developed Welfare States*. (New York: Cambridge University Press, 2002).

Duane Swank and Sven Steinmo, “The New Political Economy of Taxation in Advanced Capitalist

Democracies," American Journal of Political Science 46 (July 2002)

Bjørn Volkerink and Jakob de Haan, "Tax Ratios: A Critical Survey," OECD Tax Policy Studies No. 5 (November 2001).

Michael D. Ward and John O'Loughlin. 2002. "Spatial Processes and Political Methodology: Introduction to the Special Issue." Political Analysis 10(3):211-216.

David E. Wildasin, "Interjurisdictional Capital Mobility: Fiscal Externality and a Corrective Subsidy," Journal of Urban Economics 25 (March 1989)

John Douglas Wilson, "Theories of Tax Competition," National Tax Journal 52 (June 1999)