An m-STAR Model of Dynamic, Endogenous Interdependence – a.k.a. Network-Behavior Coevolution – in the Social Sciences *

Jude C. Hays (jchays@uiuc.edu)
Assistant Professor of Political Science, University of Illinois

Aya Kachi (akachi2@illinois.edu)
Doctoral Candidate, Department of Political Science, University of Illinois

Robert J. Franzese, Jr. (franzese@umich.edu)
Professor of Political Science, University of Michigan

8 June 2009

ABSTRACT: *Spatial interdependence*, the interdependence of outcomes across units, is theoretically and substantively ubiquitous and central across the social sciences. The empirical clustering or correlation of outcomes on some dimension(s), *spatial association*, is also obvious in most contexts. However, outcomes may exhibit spatial association for three distinct reasons. First, units may be responding similarly to similar exposure to similar exogenous internal/domestic or external/foreign stimuli (*common exposure*); second, units' responses may depend on others' responses (*contagion*). A third possibility arises when the putative outcome affects the variable along which clustering occurs (*selection*: e.g., *homophily*). Severe empirical difficulties confront the accurate estimation and distinction of these alternative sources of spatial association. After brief review of our previous work on specification, estimation, testing, and interpretation of the *spatial* and *spatiotemporal autoregressive* (*SAR* and *STAR*) models, which reflect interdependence directly and so can address *Galton's Problem* of distinguishing common exposure from contagion as alternative substantive sources of observed spatial association, this paper extends those analyses, proposing to apply the *multiparametric spatiotemporal autoregressive* (*m-STAR*) model as a simple approach to estimating jointly the pattern of connectivity and the strength of contagion by that pattern, including the case where connectivity is endogenous to the dependent variable (i.e., selection). As before, we stress substantively-theoretically guided (i.e., structural) specifications that can support analyses of estimated spatiotemporal responses to stochastic or covariate shocks and that can distinguish the possible sources of spatial association, now three: common exposure, contagion, and selection. In addition to discussing estimation of m-STAR models, this paper compares the approach to extant longitudinal-network strategies [this work still pending as yet], and suggests how to calculate, interpret, and present the dynamic, endogenous coevolution of network structure and of contagion and common-exposure effects that emerges from such a system of nonlinear endogenous equations. We illustrate this approach to dynamic, endogenous interdependence—which parallels models of network and behavior coevolution in the dynamic or longitudinal networks literature—with an empirical application attempting to disentangle the roles of economic interdependence, correlated external and internal stimuli, and EU membership in shaping labor-market policies in recent years.

* This research is supported in part by NSF grant #0318045. We thank Chris Achen, Jim Alt, Kenichi Ariga, Klaus Armingeon, Neal Beck, Jake Bowers, Jim Caporaso, Kerwin Charles, Bryce Corrigan, Tom Cusack, David Darmofal, Jakob de Haan, Scott De Marchi, John Dinardo, Zach Elkins, James Fowler, John Freeman, Fabrizio Gilardi, Kristian Gleditsch, Mark Hallerberg, John Jackson, Jonathan Katz, Mark Kayser, Achim Kemmerling, Gary King, Hasan Kirmanoglu, Herbert Kitschelt, Jim Kuklinski, Don Lacombe, Chris Lawrence, Tse-Min Lin, Xiaobo Lu, Scott McClurg, Walter Mebane, Covadonga Meseguer, Michael Peress, Thomas Pluemper, David Prosperi, Dennis Quinn, Megan Reif, Frances Rosenbluth, Ken Scheve, Phil Schrodt, Chuck Shipan, Beth Simmons, David Siroky, John D. Stephens, Duane Swank, Wendy Tam-Cho, Vera Troeger, Craig Volden, Michael Ward, Greg Wawro, and Erik Wibbels – and the many others we have surely forgotten – for useful comments on this and/or other work in our broader project on spatial-econometric models in political science. All remaining errors are ours alone.

KEYWORDS: Interdependence, Network Dependence, Spatial Lag, Spatial Econometrics.

I. <u>INTRODUCTION</u>: The Substantive and Theoretical Ubiquity and Centrality of Spatial Interdependence, Its Mechanisms, and a General Theoretical Model

Social-scientific interest in and applications of spatial-econometric modeling have burgeoned lately, due partly to advances in theory that imply interdependence and in methodology to address it; partly to global substantive developments that have raised perceptions and attention to interconnectivity, at all levels, from micro/personal to macro/international; and partly to advances in technology for obtaining and working with spatial data. This is a welcome development because the dependence of outcomes in some units on outcomes in others, *spatial interdependence*, is substantively ubiquitous and theoretically central across the political and other social sciences.

Perhaps the most-extensive classical and current political-science interest in spatial interdependence, dating from the 1950s and still booming, surrounds intergovernmental diffusion of policies among U.S. States.[1] Similar policy-diffusion research has emerged more-recently in comparative studies, but the closer parallel in classical and current comparative and international politics research regards institutional/regime diffusion, which dates at least to Dahl's (1971) *Polyarchy* and is much invigorated since the fall of the Soviet Union and Starr's (1991) "Democratic Dominoes" and Huntington's (1991) *Third Wave*.

The topical range of substantively important spatial-interdependence extends well beyond such inter-governmental diffusion, however, spanning all of political science. Inside democratic legislatures, representatives' votes depend on others' (expected) votes, and, in electoral studies, citizens' votes, election outcomes, or candidate qualities, strategies, or contributions in some contests depend on those in others. In micro-behavioral work, too, much of the longstanding and recently surging interest in contextual/*neighborhood* effects surrounds effects on respondents' behaviors or opinions of aggregates of others' (e.g., those of his/her community or social network). Contagion or diffusion in ideology, or social-movements, or national identity have also been explored. In comparative and international political economy, too, interdependence is often substantively large and central. Many stress cross-national diffusion as a force behind recent economic liberalizations, for instance. More broadly, globalization, i.e., international economic integration, arguably today's most-

[1] The ensuing list of topics and disciplines corresponds to literature searches for applied work under *contagion*, *spatial interdependence*, or *network dependence*. A web appendix provides, among other things, full citation to these (*many*) works, with some (little) annotation, topically organized in the order presented here in the text: www.umich.edu/~franzese/Publications.html.

notable (and indisputably most-noted) political-economic phenomenon, implies strategic and nonstrategic interdependence of domestic politics, policymakers, and policies. Likewise, the probability and outcomes of coups, riots, civil wars, and revolutions in one unit depend on those in others. Terrorist origins and targets also manifest spatial patterns. Simply put, the interdependence of states' actions defines the subfield of international relations.

In fact, interdependence of outcomes across units could serve as reasonable definition of *social* science more generally and broadly. Interdependence is indeed studied prominently in geographical and environmental sciences, in regional, urban, and real-estate economics, in medicine, public health, epidemiology, and criminology, and, in its related guise as network-dependence, in medicine, health, and epidemiology again, in education, and, of course, in social-network analysis. Topics include, to name just a few, interdependence in technology, marketing, and firm strategies; in macroeconomic performance; in microeconomic utilities; in violence and crime; and network dependence in obesity, fertility, birthweight, child development and poverty; in marriage; in right-wing extremism; in (sub)national identity; in women's ordainment; and in academic citations, placements, and co-authoring.

In short, as *Tobler's Law* (1970) aptly sums: "Everything is related to everything else, but near things are more related than distant things." Furthermore, as Beck et al.'s (2006) pithy title reminds in corollary: "Space is more than Geography." I.e., the substantive content of the proximity in *Tobler's Law*, and so the pathways along which interdependence between units may operate, extend well beyond physical distance, contact, and contiguity (as the examples above attest). Long literatures in sociology, regional science, geography, have elaborated from those disciplinary perspectives the multifarious mechanisms by which contagion may arise. Simmons et al. (2005, 2006) offer a list for international relations–*coercion*, *competition*, *learning*, and *emulation*–that has been influential in political science.[2]

In fact, strategic interdependence arises any time some unit(s)'s actions affect the marginal utility of other(s)'s actions.[3] Given such externalities, i's utility depends on both its choice/outcome and that of j. In environmental policy, for instance, domestic welfare (or

[2] For fuller, closer match to prior traditions, add *cooperation* and *externality* to *competition*, merge *learning* and *emulation*, and add *relocation diffusion*—direct movement of some parts of units i into other units j, such as by human migration or disease contagion (Haegerstrand 1970).

[3] Manski (2000) shows such externalities could arise in formal microeconomic models from interactions, expectations, or preferences. Akerlof (1997), Glaeser et al. (2000, 2003), Brock & Durlauf (2001), e.g., provide further examples and reviews. Non-strategic interdependence could arise even without such externalities.

net political-economic benefits to policymakers) in each country will depend on the policies of both countries due to environmental spillovers (e.g., of pollution) and economic spillovers (e.g., in regulatory costs). Optimizing behavior will yield best-response functions of *i*'s optimal policies as a function of *j*'s and *vice versa*. In this framework, moreover, positive externalities create *free-rider* incentives, which induce policies to move in opposite directions (i.e., as *strategic substitutes*), confer late-mover advantages, and make war-of-attrition (strategic delay or inaction) dynamics likely. Conversely, negative externalities create strategic complementarity, with policies moving in the same direction, yielding early-mover advantages and *competitive races* (to the bottom, top, or elsewhere).

Formally, following Brueckner (2003), consider two states (i,j), each with welfare (or indirect utility, V) that, due to externalities, depends on domestic *and* foreign policy (p_i,p_j):

$$V^i \equiv V^i(p_i, p_j) \quad ; \quad V^j \equiv V^j(p_j, p_i) \qquad (1).$$

As i chooses p_i to maximize its welfare, this affects j's optimal policy-choice, and *vice versa*. We can express such *strategic interdependence* between i and j as *best-response functions*, giving i's optimal policy, p_i^*, as a function of j's policy:

$$p_i^* = \text{Argmax}_{\text{p}_\text{i}} V^i(p_i, p_j) \equiv R(p_j) \quad ; \quad p_j^* = \text{Argmax}_{\text{p}_\text{j}} V^j(p_j, p_i) \equiv R(p_i) \qquad (2).$$

The signs of the response-function slopes determine whether *competitive-race* or *free-rider* dynamics occur; they depend on these ratios of second cross-partials:

$$\frac{\partial p_i^*}{\partial p_j} = -V^i_{p_i p_j} \,/\, V^i_{p_i p_i} \quad ; \quad \frac{\partial p_j^*}{\partial p_i} = -V^j_{p_j p_i} \,/\, V^j_{p_j p_j} \qquad (3).$$

If governments are maximizing, the denominators are negative, so, if $V^{i,j}_{p_i p_j} > 0$, policies are *strategic complements*: reaction-functions slope upward. If $V^{i,j}_{p_i p_j} < 0$, reaction functions slope downward: policies are *strategic substitutes*. If $V^{i,j}_{p_i p_j} = 0$, best-response functions are flat: strategic interdependence does not materialize. Interestingly, *negative* externalities induce strategic-*complement* policy-interdependence (i.e., positive feedback), and *positive* externalities induce strategic-*substitute* (i.e., negative) interdependence.

In our empirical application: active-labor-market (ALM) policies, assuming effectiveness, have positive employment externalities and diminishing returns, so *free-rider* dynamics should arise. Such strategic contexts also create first-mover disadvantages—those spending earlier bear larger portions of the costs of reducing unemployment—and so potential for war-of-attrition dynamics that would delay action and push equilibrium ALM spending of i

and j lower still. Do cross-border positive employment externalities of ALM policies exist; and, if so, are they sufficiently strong to induce fiscal free-riding in ALM policy? Labor-market outcomes and policies exhibit obvious spatiotemporal patterns within and across the developed democracies, and among European Union member-states especially. We have shown elsewhere (Franzese & Hays 2006c) that EU member-states' ALM policies exhibit significant interdependence along borders, a pattern possibly indicative of appreciable cross-border spillovers in labor-market outcomes inducing strategic interdependence among these political economies in labor-market policies. However, these countries also faced common or very similar exogenous-external conditions and internal trends, which would likewise tend to generate spatial patterns in the domestic policy-responses, even without interdependence. Moreover, EU membership itself likely entails both some common external stimuli and some strategic interdependencies relevant to labor-market policy. Finally, labor-market policies themselves may shape the patterns of economic exchange by which some of the policy interdependencies arise. I.e., the policies of interest may also shape the patterns of connectivity by which foreign labor-market policies affect domestic ones, a complex sort of endogeneity known as *selection* in the dynamic-networks literature.

In summary, spatial interdependence is theoretically and substantively ubiquitous and central across the social sciences, and ALM policy is likely no exception. The empirical clustering or correlation of outcomes on some dimension(s), spatial association, is also obvious in most contexts, including ALM policy. However, outcomes may exhibit spatial association for three distinct reasons. First, units may respond similarly to similar exposure to similar exogenous internal/domestic or external/foreign stimuli (*common exposure*); second, units' responses may depend on others' responses (*contagion*). We may find states' adoptions of some ALM policy-stance, for example, to cluster geographically or along other dimensions of proximity, e.g., bilateral trade-volume, because states that are proximate on that dimension experience similar exogenous domestic or foreign political-economic stimuli or because each state's ALM-policy decisions depend on what ALM policies other states proximate in this way implement. A third possibility arises when the putative outcome affects the variable along which clustering occurs (*selection*). States ALM policies might also cluster according to some variable on which we observe their proximity (bilateral trade volume) because their ALM policies affect that variable (here: spur trade between them).

Franzese & Hays discussed elsewhere (2003, 2004ab, 2006abc, 2007abcd, 2008abc) the severe empirical-methodological challenges in estimating interdependence and contagion distinctly and well (a.k.a, *Galton's Problem*[4]). Section II of this paper briefly reviews that work on specification, estimation, testing, and interpretation of *spatial* and *spatiotemporal autoregressive* (*SAR* and *STAR*) models, which reflect interdependence directly and which therefore are capable of distinguishing common exposure from contagion as alternative substantive sources of observed spatial association. There we showed that the relative and absolute accuracy and power with which the empirical-model specification reflects the patterns of interdependence on the hand and the exogenous internal and external stimuli on the other are of first-order importance in drawing such distinctions. This leads naturally to the extensions offered in Section III, where we propose applying the *multiparametric spatiotemporal autoregressive* (*m-STAR*) model as a simple means of estimating the pattern of connectivity jointly with the strength of contagion by that pattern, including the case where connectivity is endogenous to the dependent variable (i.e., *selection*; e.g., *homophily*). We again emphasize substantively-theoretically guided (i.e., structural) specifications that can support analyses of estimated spatiotemporal responses to stochastic or covariate shocks and that can distinguish the possible sources of spatial association, now three: common exposure, contagion, and selection. As before, these processes will typically look much alike empirically, so the relative omission or inadequacy in the empirical model and estimates of any one part will bias inferences in favor of others similar to it. Accordingly, valid inferences regarding any generally require empirical modeling that specifies and estimates all three processes well. Section IV offers some simulation evidence of the superior small-sample performance of full-information spatial maximum-likelihood (S-ML) estimates of m-STAR models in these contexts to that of a naïve estimator applying least-squares to a linear regression including multiple spatial-lags and to that of a blind estimator that applies least-squares omitting spatial lags. Section V illustrates this approach to dynamic, endogenous interdependence—which parallels models of *network/behavior coevolution* in the dynamic or longitudinal networks literature—with an empirical application attempting to disentangle the roles of economic interdependence, correlated external and internal stimuli, and EU membership in shaping labor-market policies in recent years, emphasizing interpretation and presentation of the estimated coevolutionary spatiotemporal dynamics.

II. Spatial & Spatiotemporal Models of Interdependence: Specification, Estimation,

[4] The web appendix contains, *inter alia*, brief intellectual-historical background to the label.

Interpretation, Presentation

To reflect *interdependence* across units of outcomes directly, empirical models should specify outcomes in units i and j as affecting each other. We suggested elsewhere (2004a, 2006a, 2007bc, 2008ab) the following such generic model of modern, open-economy, context-conditional political-economy, for example:

$$y_{it} = \rho \sum_{j \neq i} w_{ij} y_{jt} + \phi y_{i,t-1} + \boldsymbol{\beta}'_d \mathbf{d}_{it} + \boldsymbol{\beta}'_s \mathbf{s}_t + \boldsymbol{\beta}'_{sd} \left(\mathbf{d}_{it} \odot \mathbf{s}_t \right) + \varepsilon_{it} \quad (4).[5]$$

y_{jt} is the outcome in another ($j \neq i$) unit, which in some manner (given by ρw_{ij}) directly affects the outcome in unit i. The w_{ij} reflect the *relative* connectivity from j to i, and ρ reflects the overall *strength of dependence* of the outcome in i on the outcomes in the other ($j \neq i$) units, as weighted by w_{ij}. Substantively for ALM-policy interdependence, e.g., the w_{ij}, could gauge the sizes, trade, geographic contiguity, or EU comembership of i's and j's political economies. The other right-hand-side factors reflect the non-interdependence components: unit-level/domestic factors $\mathbf{d}_{it}$ (e.g., election-year indicators, government partisanship), exogenous-external/contextual factors $\mathbf{s}_{it}$ (e.g., technology, oil prices; merely for contrast, assume these common across units: $\mathbf{s}_t$), and context-conditional factors $\mathbf{d}_{it} \odot \mathbf{s}_t$ (i.e., the interactions of the former with the latter). The ε_{it} are *i.i.d.* stochastic terms.[6]

Distinguishing spatial (or network) interdependence from non-dependence sources of spatial association is the essence of *Galton's Problem*. A third potential source of spatial correlation, to be introduced later, is that the relative connectivity from j to i, that is, the w_{ij}, may depend on the outcome(s) in i (and/or j). As we summarize below (from Franzese & Hays 2003, 2004ab, 2006b, 2007abcd, 2008abc), obtaining *good* (unbiased, consistent, and efficient) parameter and certainty estimates in such models is *not* straightforward.[7] The first and prime consideration in weighing these alternatives and estimating the role of the corresponding aspects of (4) are the theoretical and empirical precision and explanatory power, relatively and absolutely, of the spatial and non-interdependence parts of the model. To elaborate: the relative and absolute accuracy and power with which the spatial weights,

[5] The $\odot$ here indicates element-by-element multiplication (i.e., Hadamard product). The model is merely heuristic, intended to encompass common classes of argument in C&IPE.

[6] One could also allow further spatial error-correlation and address it by FGLS or PCSE, or in the likelihood, but optimal will be to model interdependence and correlation in the first moment insofar as possible.

[7] Some might suggest starting with nonspatial models and adding spatial aspects as data demand, but tests that can distinguish interdependence from other potential sources of residual spatial-correlation in non-spatial models are weak (Anselin 2006; Franzese & Hays 2008b; Hendry 2006; but cf. Florax et al. 2003, 2006).

w_{ij}, reflect and offer leverage upon the *interdependence* mechanisms actually operating and with which the exogenous domestic, external, and/or context-conditional parts reflect and gain leverage upon the alternatives are crucial to the attempt to distinguish and evaluate their strength empirically. The two mechanisms produce similar effects, so inadequacies or omissions in specifying the one tend, intuitively, to induce overestimates of the other's role.

Secondarily,[8] even with the *interdependence* and the alternative *common-shock* mechanisms modeled perfectly, the spatial-lag regressor(s) will be endogenous (i.e., covary with ε), so estimates of ρ will suffer simultaneity biases. Furthermore, as with the primary concern of relative omitted-variable or misspecification bias, these simultaneity biases in estimated strength of interdependence (usually overestimation) generally induce biases in the opposite direction (underestimation) regarding the role of common shocks. Therefore, researchers who emphasize unit-level/domestic, exogenous-external, or context-conditional processes to the exclusion or relative neglect of interdependence will tend to get empirical results biased toward the former and against the latter sorts of explanations. Conversely, researchers stressing interdependence to the relative neglect of domestic/unit or exogenous-contextual considerations or who fail to account sufficiently the endogeneity of spatial lags will tend to suffer the opposite biases: underestimating the role of exogenous domestic, external, or context-conditional factors and overestimating that of interdependence.

Most empirical studies in comparative and international political economy (C&IPE) where interdependence may arise, especially those in the policy diffusion, globalization, tax-competition, and policy-autonomy literatures, analyze time-series cross-sections (TSCS). In such contexts, employing spatial and temporal lags to specify both temporal and spatial dependence directly in a spatiotemporal autoregressive (STAR) model is often desirable:[9]

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \phi\mathbf{M}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (5).$$

The dependent variable, $\mathbf{y}$, is an $NT{\times}1$ vector of cross sections stacked by periods (i.e., the N first-period observations, the next N, up through N in period T).[10] ρ is the previously

[8] Simulations (Franzese & Hays 2004a, 2006b, 2007cd) show the omitted-variable/misspecification biases of omission/relative-neglect of interdependence typically far exceed the simultaneity biases of failing to redress adequately the spatial-lag endogeneity, although the latter grow appreciable as interdependence strengthens.

[9] Anselin (2002, 2006) distinguishes spatial statistics and spatial econometrics in methodological approach as the former being more data-driven and tending toward treating spatial correlation as nuisance and the latter wedded more to theoretically structured models of interdependence. The web appendix offers fuller discussion of this subtle but key distinction. In these terms, our approach is a decidedly spatial-econometric one.

[10] Nonrectangular or missing data are manageable, but rectangularity is assumed here for expository ease.

described spatial-autoregressive coefficient, and $\mathbf{W}_{NT}$ is an $NT \times NT$ block-diagonal spatial-weighting matrix.[11] Each of the T $N \times N$ weights matrices, $\mathbf{W}_N^t$, on the block-diagonal have elements $w_{ij}(t)$ reflecting the relative connectivity from unit j to i that period.[12] Thus, for each observation, y_{it}, the *spatial lag*, $\mathbf{Wy}$, gives a weighted sum of the y_{jt}, with weights $w_{ij}(t)$ being direct and straightforward reflection of the dependence of each unit i's outcome on others'. $\mathbf{M}$ is an $NT \times NT$ matrix with ones on the minor diagonal, i.e., at coordinates $(N+1,1)$, $(N+2,2)$, …, $(NT,NT\text{-}1)$, and zeros elsewhere. $\mathbf{My}$ is thus a standard (first-order) temporal-lag;[13] ϕ is its coefficient. $\mathbf{X}$ contains NT observations on k independent variables; $\boldsymbol{\beta}$ is its $k \times 1$ vector of coefficients, and $\boldsymbol{\varepsilon}$ is an $NT \times 1$ vector of *i.i.d.* stochastic components.[14]

Franzese & Hays (2004a, 2006b, 2007cd, 2008b) explored analytically and by simulation the properties of four estimators for such models: non-spatial least-squares (i.e., regression omitting the spatial component as is common in most extant research: OLS), spatial OLS (i.e., OLS estimation of models like (5), common in diffusion studies and becoming so in globalization/tax-competition ones: S-OLS), instrumental variables (e.g., spatial 2SLS or S-2SLS), and spatial maximum-likelihood (S-ML). Both OLS and spatial OLS produce biased and inconsistent estimates, OLS due to the omitted-variable bias and spatial OLS because the spatial lag is endogenous and so induces simultaneity bias. We can view these biases as reflecting the terms of *Galton's Problem.* On one hand, by omitting the spatial lag that would reflect the interdependence, OLS coefficient-estimates will suffer omitted-variable biases—familiarly: $\mathbf{F}\boldsymbol{\beta}$, where $\mathbf{F}$ is the matrix of coefficients obtained by regressing the omitted on the included variables and $\boldsymbol{\beta}$ is the vector of (true) coefficients on the omitted variables.[15] In this case, the omitted-variable bias (OVB) is:

$$\text{OVB}\left[\hat{\boldsymbol{\beta}}_{\text{OLS}} \quad \hat{\phi}_{\text{OLS}}\right]' = \rho \times \left(\mathbf{Q}_1'\mathbf{Q}_1\right)^{-1}\mathbf{Q}_1'\mathbf{Wy}, \text{ where } \mathbf{Q}_1 \equiv \left[\mathbf{X} \quad \mathbf{My}\right]' \qquad (6).$$

$\hat{\rho}_{\text{OLS}} \equiv 0$, of course, which is biased by $-\rho$. Thus, insofar as the spatial lag covaries with the non-spatial regressors—which is (i) highly likely if domestic conditions correlate spatially,

[11] $\mathbf{W}_{NT}$ is block-diagonal assuming no cross-temporal spatial interdependence. Non-zero off-diagonal blocks are possible and manageable, but perhaps unlikely controlling for time lags and contemporaneous spatial-lags.

[12] If the pattern of connectivity is time-invariant, then $\mathbf{W}_{NT}$ can be expressed as the Kronecker product of a $T \times T$ identity matrix and the constant $N \times N$ weights-matrix, $\mathbf{I}_T \otimes \mathbf{W}_N$.

[13] Higher-order time-dynamics would simply add further properly configured weights matrices.

[14] Again, alternative distributions of $\boldsymbol{\varepsilon}$ are possible but add complication without illumination.

[15] Estimates of limited- or qualitative-dependent-variable models, like logit or probit, which exclude relevant spatial lags will suffer analogous omitted-variable biases.

(*ii*) certain for exactly common exogenous-external shocks, and (*iii*), given non-zero spatial correlation from any source, certain for the time lag also—OLS will overestimate domestic, exogenous-external, or context-conditional effects, including the temporal adjustment-rate, while ignoring interdependence. On the other hand, including spatial lags in models for OLS estimation raises inherent endogeneity biases. Spatial lag, $\mathbf{Wy}$, covaries with the residual, $\boldsymbol{\varepsilon}$, making S-OLS estimates inconsistent, because it is a weighted average of outcomes in other units and so places some observations' left-hand sides on the right-hand sides of others: textbook simultaneity. In simplest terms by example: Germany causes France, but France also causes Germany. These asymptotic simultaneity biases (SB) are:

$$\text{SB}\left[\hat{\rho} \quad \hat{\phi} \quad \hat{\boldsymbol{\beta}}\right]' = \left(\mathbf{Q}'\mathbf{Q}\right)^{-1}\mathbf{Q}'\boldsymbol{\varepsilon}, \text{ where } \mathbf{Q} \equiv \left[\mathbf{Wy} \quad \mathbf{My} \quad \mathbf{X}\right] \qquad (7).$$

In the case where $\mathbf{X}$ contains just one exogenous explanator, $\mathbf{x}$, these biases are:

$$\text{SB}\begin{bmatrix}\hat{\rho}\\ \hat{\phi}\\ \hat{\beta}\end{bmatrix} = \frac{1}{|\boldsymbol{\Psi}|}\begin{bmatrix}\text{Cov}\left(\mathbf{Wy},\varepsilon\right)\times\text{Var}\left(\mathbf{My}\right)\times\text{Var}\left(\mathbf{x}\right)\\ -\text{Cov}\left(\mathbf{Wy},\varepsilon\right)\times\text{Cov}\left(\mathbf{Wy},\mathbf{My}\right)\times\text{Var}\left(\mathbf{x}\right)\\ -\text{Cov}\left(\mathbf{Wy},\varepsilon\right)\times\text{Cov}\left(\mathbf{Wy},\mathbf{x}\right)\times\text{Var}\left(\mathbf{My}\right)\end{bmatrix}, \text{ where } \boldsymbol{\Psi} = \text{plim}\left(\frac{\mathbf{Q}'\mathbf{Q}}{n}\right) \qquad (8).$$

With positive interdependence and positive covariance of the spatial-lag with the exogenous regressors, a likely common case, one overestimates the interdependence-strength, $\hat{\rho}$, and correspondingly underestimates temporal dependence, $\hat{\phi}$, and exogenous effects, $\hat{\beta}$.

In sum, *Galton's Problem* implies that empirical analyses that ignore substantively appreciable interdependence will also thereby tend to overestimate the importance of non-spatial factors, with the effect of factors that correlate spatially the most, in pattern most similar to $\mathbf{W}$, being most overestimated. On the other hand, simple controls for spatial-lag processes (or studying them qualitatively) will suffer simultaneity biases, usually in the opposite direction, exaggerating interdependence and understating unit-level/domestic, exogeneous-external, and context-conditional effects. Again, those factors that correlate most with the interdependence pattern will have the most severe induced deflation biases. These conclusions hold as a matter of degree as well; insofar as the non-spatial components of the model are inadequately specified and measured relative to interdependence aspects, the latter will be privileged and the former disadvantaged, and *vice versa.* Accurate and powerful specification of $\mathbf{W}$ is therefore of crucial empirical, theoretical, and substantive importance, obviously for those interested in interdependence, but also for those primarily interested in domestic/unit-level, exogenous-external/contextual, or context-conditional factors. Conversely, optimal specification of the unit-level/domestic, contextual/exogenous-external, and context-conditional non-spatial components is of equally crucial importance to those interested in gauging the importance of interdependence.

Our simulations (Franzese & Hays 2004a, 2006b, 2007cd) showed the omitted-variable biases of OLS are almost always worse and often far, far worse than S-OLS' simultaneity biases. In fact, S-OLS may perform adequately for mild interdependence strengths ($\rho\Sigma_j w_{ij} \lesssim .25$), although standard-error accuracy can be problematic, and in a manner for which PCSE (Beck & Katz 1995, 1996) will not compensate. S-OLS' simultaneity biases grow sizable as interdependence strengthens, however, rendering use of a consistent estimator, such as S-2SLS or S-ML, highly advisable. Choosing which consistent estimator seems of secondary importance in bias, efficiency, and standard-error-accuracy terms. Since S-ML proved close to weakly dominant,[16] we introduce only it here.[17,18]

The conditional likelihood function for the spatiotemporal-lag model,[19] which assumes the first observations non-stochastic, is a straightforward extension of the standard spatial-lag likelihood function, which in turn adds only one mathematically and conceptually small complication to the likelihood function for the standard linear-normal model (OLS). To see this, start by rewriting the spatial-lag model with the stochastic component on the left:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \Rightarrow \boldsymbol{\varepsilon} = \left(\mathbf{I} - \rho\mathbf{W}\right)\mathbf{y} - \mathbf{X}\boldsymbol{\beta} \equiv \mathbf{A}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} \qquad (9),$$

where $\mathbf{X}$ now includes $\mathbf{My}$, the time-lag of $\mathbf{y}$, as its first column, and $\boldsymbol{\beta}$ includes ϕ as its first

[16] See Franzese & Hays (2007b, 2008b) regarding S-ML estimation; they correct some misleading conclusions from our earlier work on S-ML, stemming from a coding error. (An error in LeSage's original code called the wrong element of the estimated variance-covariance matrix as standard errors of the spatial-lag coefficient.) The instrumental-variables (IV), two-stage-least-squares (2SLS), generalized-method-of-moments (GMM) family of estimators relies on the spatial structure of the data to instrument for the endogenous spatial lag. Assuming no *cross-spatial endogeneity* (our term for y's in some units causing $\mathbf{x}$'s in others), $\mathbf{WX}$ are ideal instruments by construction. Cross-spatial endogeneity may seem unlikely, until one realizes that vertical ties y_i to $\mathbf{y}_j$ (interdependence) and horizontal ties from $\mathbf{y}_j$ to $\mathbf{x}_j$ (typical simultaneity) combine to give the offending diagonals from y_i to $\mathbf{x}_j$. However, S-GMM should improve upon S-2SLS primary weakness in efficiency, so it may compare more favorably to S-ML. Estimation by instrumentation may also prove more robust in some ways—e.g., to non-normal distributional issues—than S-ML. We have not yet explored these possibilities.

[17] Initially, we used J.P. LeSage's MatLab™ code to estimate our spatial models, having found third-party contributed Stata™ code for spatial analysis untrustworthy and/or extremely computer-time intensive. We have since written MatLab™ code to implement all, and Stata™ code to implement many, of our suggestions. For code, plus Excel™ spreadsheets useful as templates for interpretation and presentation: https://netfiles.uiuc.edu/jchays/www/page.html and http://www.umich.edu/~franzese/Publications.html.

[18] Franzese & Hays (2008d), Hays & Kachi (2008), and Hays (2009) introduce for political science empirical models of spatial interdependence in limited and qualitative dependent-variables.

[19] Derivation of likelihoods for spatiotemporal-lag models is due to Elhorst (2001, 2003, 2005).

row.[20] Assuming *i.i.d.* normality, the likelihood function for $\boldsymbol{\varepsilon}$ is the typical linear-normal:

$$L(\boldsymbol{\varepsilon}) = \left(\frac{1}{\sigma^2 2\pi}\right)^{\frac{NT}{2}} \exp\left(-\frac{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{2\sigma^2}\right) \tag{10},$$

which will produce a likelihood in terms of $\mathbf{y}$ as follows:

$$L(\mathbf{y}) = \mid \mathbf{A} \mid \left(\frac{1}{\sigma^2 2\pi}\right)^{\frac{NT}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(\mathbf{Ay} - \mathbf{X\beta}\right)'\left(\mathbf{Ay} - \mathbf{X\beta}\right)\right) \tag{11}.$$

This resembles the typical linear-normal likelihood, except the transformation from $\boldsymbol{\varepsilon}$ to $\mathbf{y}$ is not by the usual factor, 1, but by $|\mathbf{A}|=|\mathbf{I}\text{-}\rho\mathbf{W}|$.[21] Written in $(N\times1)$ vector notation, the spatiotemporal-model conditional-likelihood is mostly conveniently separable into parts:

$$\begin{gathered} \text{Log } f_{\mathbf{y}_t,\mathbf{y}_{t-1},\dots,\mathbf{y}_2|\mathbf{y}_1} = -\frac{1}{2}N\left(T-1\right)\log\left(2\pi\sigma^2\right) + \left(T-1\right)\log\left|\mathbf{A}\right| - \frac{1}{2\sigma^2}\sum_{t=2}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t \\ \text{where } \boldsymbol{\varepsilon}_t = \left(\mathbf{I}_N - \phi\mathbf{I}_N\right)\mathbf{y}_t - \rho\mathbf{W}_N\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta}. \end{gathered} \tag{12}.$$

The unconditional (exact) likelihood function, which retains the first time-period observations as non-predetermined, is more complex (Elhorst 2005):

$$\begin{aligned} &\text{Log } f_{\mathbf{y}_t,\dots,\mathbf{y}_1} = -\tfrac{1}{2}NT\times\log(2\pi\sigma^2) + \tfrac{1}{2}\sum_{i=1}^{N}\log\left((1-\rho\omega_i)^2 - \phi^2\right) + (T-1)\log\mid\mathbf{A}\mid \\ &-\tfrac{1}{2\sigma^2}\sum_{t=2}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t - \tfrac{1}{2\sigma^2}\boldsymbol{\varepsilon}_1'(\mathbf{A}-\phi\mathbf{I}_N)'^{-1}[\mathbf{A}'\mathbf{A} - \phi^2\mathbf{A}'\mathbf{A}^{-1}(\mathbf{A}'\mathbf{A}^{-1})']^{-1}(\mathbf{A}-\phi\mathbf{I}_N)^{-1}\boldsymbol{\varepsilon}_1 \\ &\text{where } \boldsymbol{\varepsilon}_1 = (\mathbf{I}_N - \phi\mathbf{I}_N)\mathbf{y}_1 - \rho\mathbf{W}_N\mathbf{y}_1 - \mathbf{x}_1'\boldsymbol{\beta}. \end{aligned} \tag{13}.$$

With large T, the first observation contributes little to the total likelihood, so the simpler conditional likelihood can serve adequately.

One easy way to ease or even erase the simultaneity problem with S-OLS is to lag temporally the spatial lag (Beck et al. 2006). Insofar as time-lagging the spatial lag renders it pre-determined—i.e., to the extent interdependence does not incur instantaneously, where *instantaneous* means *within an observation period, as measured, given the model*—S-OLS' bias disappears asymptotically. Formally, the STAR model with time-lagged spatial-lag is:

$$\mathbf{y}_t = \eta\mathbf{W}\mathbf{y}_{t-1} + \phi\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \tag{14}.$$

Elhorst (2001:126-30) derives the unconditional log-likelihood for this model as:

$$\begin{gathered} \text{Log } f_{\mathbf{y}_t,\dots,\mathbf{y}_1} = -\tfrac{1}{2}NT\times\log(2\pi\sigma^2) + \tfrac{1}{2}\sum_{i=1}^{N}\log\left(1-(\phi+\eta\omega_i)^2\right) - \tfrac{1}{2\sigma^2}\sum_{t=2}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t \\ -\tfrac{1}{2\sigma^2}\boldsymbol{\varepsilon}_1'(\mathbf{I}-\mathbf{B})'^{-1}(\mathbf{I}-\mathbf{BB}')^{-1}(\mathbf{I}-\mathbf{B})^{-1}\boldsymbol{\varepsilon}_1 \end{gathered} \tag{15},$$

[20] N.b., although $\mathbf{Wy}$ complicates the conditional likelihood in terms of $\mathbf{y}$ (see note 21), $\mathbf{My}$ does not.

[21] This difference complicates estimation somewhat in that the determinant $|\mathbf{A}|$ involves ρ, and so requires recalculation at each iteration of the likelihood-maximization routine.

where $\boldsymbol{\varepsilon}_1 = \mathbf{y}_1 - (\phi + \eta \mathbf{W}_N)\mathbf{y}_1 - \mathbf{X}_1\boldsymbol{\beta}$, $\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \eta \mathbf{W}_N \mathbf{y}_{t-1} - \phi \mathbf{y}_{t-1} - \mathbf{X}_t\boldsymbol{\beta}$, and $\mathbf{B} = \phi \mathbf{I}_N + \eta \mathbf{W}_N$. Note that the second and fourth terms in (15) bias OLS estimation of (14). Asymptotically ($T \to \infty$), the contribution of these terms to the likelihood and so this bias goes to zero. In sum, if T is large, if spatial-interdependence processes operate only with a time lag and not within an observational period, if observational periodization matches that of the actual spatiotemporal dynamics, and if spatiotemporal dynamics are modeled well enough for these conditions not to become violated through measurement error or misspecification leaving some time-lagged interdependence to bleed into the contemporaneous, OLS with a time-lagged spatial-lag on the RHS is an effective estimation strategy. However, even in this best case, OLS with time-lagged spatial-lags only yields unbiased estimates if first observations is non-stochastic (i.e., with initial conditions fixed across repeated samples).

Testing for remaining temporal *and* spatial correlation in OLS residuals is possible and, especially advisable if applying OLS to the time-lagged spatiotemporal-lag model. Standard Lagrange-multiplier (LM) tests for remaining temporal correlation remain valid. Following Anselin (1996), Franzese & Hays (2008b) describe several LM tests of spatial correlation that retain validity when applied to OLS estimated residuals from models containing spatial and temporal lags.[22] E.g., a standard one-directional test against spatial-lag alternative is:

$$\begin{aligned} &LM_\rho = \hat{\sigma}_\varepsilon^2 \left(\hat{\boldsymbol{\varepsilon}}' \mathbf{W} \mathbf{y} / \hat{\sigma}_\varepsilon^2 \right)^2 \left[G + R \hat{\sigma}_\varepsilon^2 \right]^{-1}, \text{ where} \\ &G = (\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}}) \text{ and } R = \mathrm{tr}[(\mathbf{W}' + \mathbf{W})\mathbf{W}] \end{aligned} \quad (16),$$

and Anselin's (1996) robust one-directional test against spatial-lag alternative is

$$LM_\rho^* = G^{-1} \hat{\sigma}_\varepsilon^2 \left(\hat{\boldsymbol{\varepsilon}}' \mathbf{W} \mathbf{y} / \hat{\sigma}_\varepsilon^2 - \hat{\boldsymbol{\varepsilon}}' \mathbf{W} \hat{\boldsymbol{\varepsilon}} / \hat{\sigma}_\varepsilon^2 \right)^2 \quad (17).$$

Lastly, regarding stationarity, the conditions and issues arising in spatiotemporally dynamic models are reminiscent of those in the more familiar solely time-dynamic models. Let ω be an eigenvalue of $\mathbf{W}$; then the spatiotemporal process is covariance stationary if:

$$\left| \phi \left(\mathbf{I} - \rho \mathbf{W} \right)^{-1} \right| < 1, \quad \text{or, equivalently, if} \quad \begin{cases} \left| \phi \right| < 1 - \rho \omega_{\max}, & \text{if } \rho \geq 0 \\ \left| \phi \right| < 1 - \rho \omega_{\min}, & \text{if } \rho < 0 \end{cases} \quad (18).$$

For instance, with positive temporal and spatial dependence and $\mathbf{W}$ row-standardized, the $\omega_{\max}$=1, so stationarity familiarly requires $\phi + \rho < 1$.

Interpretation of *effects* in empirical models with spatiotemporal interdependence, as in any model beyond the strictly linear additive-separable, involves more than simply eyeing

[22] Be sure to note the corrections posted here: http://www.umich.edu/~franzese/Publications.html.

coefficient estimates. With spatiotemporal, as with solely temporal, dynamics, coefficients on regressors give only the pre-dynamic impetuses to the outcome associated with changes of those regressors. I.e., coefficients represent only the (often inherently unobservable) pre-interdependence impetus to outcomes from each regressor. Calculation of *spatiotemporal multipliers* allows expression of the estimated dependent-variable responses across all units to shocks to covariates or to the error terms in any unit(s), accounting the spatiotemporal dynamics. These multipliers also afford estimation of the long-run, steady-state, or equilibrium[23] effect of permanent shocks.[24] We apply the delta method to derive analytically the approximate estimated asymptotic variance-covariance (standard errors) for these response-path or long-run-effect estimates; standard errors can also be simulated of course.

One calculates the cumulative, steady-state spatiotemporal effects most conveniently working with the STAR model in (N×1) vector form:

$$\mathbf{y}_t = \rho \mathbf{W} \mathbf{y}_t + \phi \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \qquad (19).$$

Set $\mathbf{y}_{t-1}$ equal to $\mathbf{y}_t$ fix exogenous RHS terms, $\mathbf{X}$ and/or $\boldsymbol{\varepsilon}$, to their hypothetical permanent post-shock levels, and solve for the long-run steady-state level of $\mathbf{y}$ (assuming stationarity):[25]

$$\begin{aligned}\mathbf{y}_t &= \rho \mathbf{W} \mathbf{y}_t + \phi \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t = \left(\rho \mathbf{W} + \phi \mathbf{I}\right) \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t = \left[\mathbf{I}_{\mathrm{N}} - \rho \mathbf{W} - \phi \mathbf{I}_{\mathrm{N}}\right]^{-1} \left(\mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right) \\ &= \begin{bmatrix} 1-\phi & -\rho w_{1,2} & \cdots & \cdots & -\rho w_{1,\mathrm{N}} \\ -\rho w_{2,1} & 1-\phi & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 1-\phi & -\rho w_{(\mathrm{N}-1),\mathrm{N}} \\ -\rho w_{\mathrm{N},1} & \cdots & \cdots & -\rho w_{\mathrm{N},(\mathrm{N}-1)} & 1-\phi \end{bmatrix}^{-1} \left(\mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right) \equiv \mathbf{S} \times \left(\mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right)\end{aligned} \qquad (20).$$

Decomposing $\boldsymbol{\varepsilon}_t = \boldsymbol{\delta} + \boldsymbol{\upsilon}_t$ with $\boldsymbol{\delta}$ fixed and $\boldsymbol{\upsilon}_t$ stochastic is conceptually useful in considering the responses across units to counterfactual shocks to outcomes in others. To offer standard errors for these estimates by the delta method,[26] first denote the i^{th} column of $\mathbf{S}$ as $\mathbf{s}_i$ and its estimate as $\hat{\mathbf{s}}_i$. The steady-state spatiotemporal equilibrium responses to a one-unit increase in the i^{th} element of $\boldsymbol{\delta}$ are then $\mathbf{s}_i$, with asymptotic approximate variance-covariance matrix:

---

[23] We use the terms *long-run*, *steady-state*, and *equilibrium* effects loosely, interchangeably, to refer to the estimated asymptotic level of outcomes $\mathbf{y}$ following a hypothetical *permanent* shock.

[24] Anselin (2003) and Franzese & Hays (2006b,2007cd, 2008bcd) discuss these multipliers more fully.

[25] In the case of time-variant $\mathbf{W}_N$, one must also fix $w_{ij}(t)$ to some desired set of values.

[26] We have used only first-order approximations. Higher orders would presumably yield greater accuracy; simulation (e.g., parametric bootstrapping) may also be advantageous.

$$\widehat{\mathbf{V}\left(\hat{\mathbf{s}}_i\right)} = \left[\frac{\partial \hat{\mathbf{s}}_i}{\partial \hat{\boldsymbol{\theta}}}\right] \widehat{\mathbf{V}\left(\hat{\boldsymbol{\theta}}\right)} \left[\frac{\partial \hat{\mathbf{s}}_i}{\partial \hat{\boldsymbol{\theta}}}\right]' \tag{21}$$

Here, $\hat{\boldsymbol{\theta}} \equiv \left[\hat{\rho} \quad \hat{\phi}\right]'$, $\left[\frac{\partial \hat{\mathbf{s}}_i}{\partial \hat{\boldsymbol{\theta}}}\right] \equiv \left[\frac{\partial \hat{\mathbf{s}}_i}{\partial \hat{\rho}} \quad \frac{\partial \hat{\mathbf{s}}_i}{\partial \hat{\phi}}\right]$, and the vectors $\left[\frac{\partial \hat{\mathbf{s}}_i}{\partial \hat{\rho}}\right]$ and $\left[\frac{\partial \hat{\mathbf{s}}_i}{\partial \hat{\phi}}\right]$ are the i^{th} columns of $\hat{\mathbf{S}}\mathbf{W}\hat{\mathbf{S}}$ and of $\hat{\mathbf{S}}\hat{\mathbf{S}}$. Similarly, the steady-state spatiotemporal responses to a one-unit increase in explanatory variable k in unit i are $\mathbf{s}_i\beta_k$, with delta-method standard-errors of

$$\widehat{\mathbf{V}\left(\hat{\mathbf{s}}_i\hat{\beta}_k\right)} = \left[\frac{\partial \hat{\mathbf{s}}_i\hat{\beta}_k}{\partial \hat{\boldsymbol{\theta}}}\right] \widehat{\mathbf{V}\left(\hat{\boldsymbol{\theta}}\right)} \left[\frac{\partial \hat{\mathbf{s}}_i\hat{\beta}_k}{\partial \hat{\boldsymbol{\theta}}}\right]' \tag{22}$$

with $\hat{\boldsymbol{\theta}} \equiv \left[\hat{\rho} \quad \hat{\phi} \quad \hat{\beta}_k\right]'$ and $\left[\frac{\partial \hat{\mathbf{s}}_i\hat{\beta}_k}{\partial \hat{\boldsymbol{\theta}}}\right] \equiv \left[\frac{\partial \hat{\mathbf{s}}_i\hat{\beta}_k}{\partial \hat{\rho}} \quad \frac{\partial \hat{\mathbf{s}}_i\hat{\beta}_k}{\partial \hat{\phi}} \quad \hat{\mathbf{s}}_i\right]$, and with $\left[\frac{\partial \hat{\mathbf{s}}_i\hat{\beta}_k}{\partial \hat{\rho}}\right]$ and $\left[\frac{\partial \hat{\mathbf{s}}_i\hat{\beta}_k}{\partial \hat{\phi}}\right]$ the i^{th} columns of $\hat{\beta}_k\hat{\mathbf{S}}\mathbf{W}\hat{\mathbf{S}}$ and $\hat{\beta}_k\hat{\mathbf{S}}\hat{\mathbf{S}}$, respectively.

One can find the spatiotemporal response path of the N×1 vector of unit outcomes, $\mathbf{y}_t$, to exogenous right-hand-side terms, $\mathbf{X}$ and $\boldsymbol{\varepsilon}$, by rearranging (19) to isolate $\mathbf{y}_t$ on the left:

$$\mathbf{y}_t = \left[\mathbf{I}_N - \rho\mathbf{W}_N\right]^{-1}\left\{\phi\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right\} = \mathbf{S}\left\{\phi\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right\} \tag{23}$$

This formula gives response-paths of all units to hypothetical shocks to $\mathbf{X}$ or $\boldsymbol{\varepsilon}$ in any unit(s) $\{j\}$, including shocks in $\{i\}$ itself/themselves, by setting $(\mathbf{X}_t\boldsymbol{\beta}+\boldsymbol{\varepsilon}_t)$ to the value(s) reflecting that hypothetical in row(s) $\{j\}$. For the marginal spatiotemporal effects (non-cumulative) or to plot the over-time path of responses to a permanent change in some $\mathbf{x}$ (cumulative), and their standard errors, working with the NT×NT matrix is easier. Redefine $\mathbf{S}$ in (20) as $\mathbf{S}\equiv[\mathbf{I}_{NT}\text{-}\rho\mathbf{W}\text{-}\phi\mathbf{M}]^{-1}$ and follow the steps just given. We calculate estimated responses like these in presenting our empirical application below.

III. The Multiparametric Spatiotemporal (m-STAR) Model

As noted above, model specifications that omit spatial lags assume zero interdependence by construction; as we have shown analytically and by simulation, these omitted-variable biases tend to inflate the estimated effects of non-interdependence model-components. For instance, most extant globalization studies, having neglected spatial lags, likely overestimated the effects of domestic and exogenous-external factors while effectively preventing globalization-induced interdependence from manifesting empirically. Conversely, standard regression estimates of models with spatial lags suffer simultaneity biases. Such models have grown more common recently among researchers interested in interdependence and have long been the norm in studies of policy-diffusion and microbehavioral contextual effects. Although our previous analyses have shown that inclusion of spatial lags in simple regression models is a vast improvement over non-spatial estimation strategies, these simultaneity biases will tend to have inflated estimated interdependence strength at the expense of domestic/unit-level, exogenous-external, and context-conditional factors. The spatial-ML approach just described effectively redresses these simultaneity issues.

Above all, most crucial to proper estimation, distinction, and weighing of the strengths of interdependence and other possible sources of spatial or network association are the relative and absolute accuracy and power with which the patterns of interconnectivity and the non-interdependence aspects of the model are specified. Accordingly, strategies to estimate **W** within models in which *unobserved* patterns of interconnections among units *affect* their choices/outcomes have long interested spatial econometricians greatly, although progress has been modest. For network analysts, contrarily, estimation of the processes generating ties in the *observed* network, as opposed to the effects on this unit's choices or outcomes of others' actions as weighted by the network, is typically the *dependent variable* of the study. Network models usually take the characteristics of units, including their actions or behaviors, as *given, exogenous* explanators of what ties, typically exclusively binary ties, will form between actors. From the network-analytic perspective, spatial-econometric attempts to parameterize and, ultimately, to endogenize w_{ij} within models of interdependent unit outcomes mirror network-analytic attempts to model the *coevolution of behavior and network*. Interdependence in typical C&IPE contexts may raise additional challenges, however, in that relative connectivity is often of degree rather than binary and, more dauntingly, that the effective connectivity may not be directly observed. Rather, quite commonly, one might observe only some covariates theorized to relate to the effective connection. In the context of interdependent ALM-policymaking, for instance, many of the theorized connections arise through inherently unobservable economic competition in labor, capital, or goods markets. We observe only trade or capital flows or other symptoms of or contributors to competition.[27] In the network-analysis tradition, Leenders (1995, 1997) and Snijders and colleagues (1997, 2001, 2005, 2007ab) have advanced furthest on this crucial

[27] Both these distinctions may reflect simplifying assumptions typical of applied network analysis more than any underlying substance of social networks. Ties in friendship networks, e.g., are in truth more of degree than binary, and we often may not observe that even as directly as by survey response gauging said closeness.

next task in empirical modeling of networks-*cum*-interdependence. We briefly review their approaches next, but then suggest another possible inroad, a much simpler, yet perhaps productive, approach: adapting the *m-STAR* model to the purpose.

In Leenders' (1995, 1997), "actors…shape their networks and, simultaneously, are influenced by the structure of the network." He terms *contagion* the effects of networks on actors' *attributes*, understood broadly to mean characteristics, actions, beliefs, policies, etc. More exactly, these are the effects of others' (*alters'*) attributes on one's own (*ego*'s), where network structure determines which alters matter and how much. Leenders terms *selection* the reverse process, in which actors' attributes shape the network. In his selection model, the equivalent of w_{ij} arise by a continuous-time Markov process—to be observed at discrete-time intervals in a dataset—where an *arc* (i.e., a binary tie) from j to i forms, w_{ij}=1, or dissolves, w_{ij}=0, at rates, λ_{0ij} and λ_{1ij}, given by some observable attribute(s) of i and/or j:

$$\lambda_{0ij} = \lambda_0 + \nu_0 d_{ij} \quad ; \quad \lambda_{1ij} = \lambda_1 + \nu_1 d_{ij} \qquad (24),$$

with d_{ij} a measure of similarity of actors i and j. Leenders' contagion is a spatial-lag model:

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (25),$$

which could extend to the standard spatiotemporal model, (5), straightforwardly. Leenders (1997) integrates these contagion and selection models thus. First, let $\mathbf{A}_t$ be the $N\times N$ matrix of current realizations of (24), $\mathbf{y}_t$ be the $N\times 1$ vector of attributes for the actors, and $\mathbf{X}_t$ the $N\times k$ matrix of exogenous explanators thereof. Leenders (1997:172) expresses $\mathbf{W}_t$ as the function $\mathbf{W}_t = W(\mathbf{A}_t)$, which could be a very useful extension toward the parameterized modeling of unobserved and potentially continuous degrees of connection as a function of observed binary *arcs* (modeled by (24)), but the function as currently implemented is just the identity. The model is then identified for estimation of $\boldsymbol{\lambda}$, $\boldsymbol{\nu}$, ρ, and $\boldsymbol{\beta}$ from $\mathbf{W}_t$ and $\mathbf{y}_t$ observed at discrete intervals $t=\{1 \ldots T\}$ by time lags and the assumption that temporal implies causal precedence and that the first observation is fixed and given (raising all the issues noted above in those regards). The combined model is then:[28]

$$\mathbf{W}_t = f(\mathbf{W}_{t-1}, \mathbf{y}_t) \quad ; \quad w_{ij,t} \equiv d_{ij,t} = | y_{i,t} - y_{j,t} | \quad ; \quad y_t = \rho_1 \mathbf{W}_t \mathbf{y}_t + \rho_2 \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \quad (26)$$

He then generates $\mathbf{W}_0$, $\mathbf{y}_0$, $\{\boldsymbol{\varepsilon}_t\}$, and $\{\mathbf{X}_t\}$ randomly, and assesses by simulation the biases entailed in estimating from data collected at intervals of increasing length (measured in

[28] Leenders (1997:173-4) actually converts (25) to a temporally dynamic model like (5) by what amounts to an error-correction model, with equilibrium y being another, constant parameter to be estimated, μ, interpreted as a societal norm for y. We have simplified to a first-order time-lag to enhance comparability in exposition.

numbers of simulation periods) and in erroneously estimating only the selection process, (24), or only the contagion process that is the last expression of (26). The text leaves unclear the exact experiments and estimation procedures, so we can interpret his results only uncertainly. He seems to find, first, that increasing granularity in the periodicity of observation generally causes attenuation bias in estimates of the selection-model parameters and inflation bias in estimates of the contagion-model parameters; second, that estimated contagion is greatly inflated when selection is unmodeled but present; and third, that estimated selection is mildly inflated when contagion is present but unmodeled.

Snijders' and colleagues'[29] approach is more elaborate. In Steglich et al. (2007), they emphasize as do we that the challenge for disentangling the sources of network association (a.k.a., spatial correlation) is threefold. One must distinguish *influence* or *contagion* (a.k.a., *interdependence*), from *selection* (e.g., *homophily*), from *social contexts* (i.e., exogenous internal and/or external conditions) because any omissions or inadequacies in modeling those distinct sources of network or spatial correlation will bias conclusions in favor of the included or better-modeled mechanisms. Then, they also stress three fundamental issues confronting such attempts: observations in discrete time-intervals of continuous-time processes, the need to control for alternative mechanisms and pathways by which observed networks and outcomes may have arisen, and the network dependence of the actors which precludes estimation by common statistical techniques, most of which assume independence.

To surmount these issues in distinguishing these alternative mechanisms, they model the coevolution of networks and behavior thus. N actors are connected as given by an observed, binary, endogenous, and time-variant connectivity matrix, $\mathbf{x}$, with elements $\mathbf{x}_{ij}(t)$—in our notation, $\mathbf{W}_N$, with elements $w_{ij}(t)$. The vector of N observed, binary behaviors, $\mathbf{z}$, at time t has elements $\mathbf{z}_i(\mathrm{t})$—in our notation, $\mathbf{y}(t)$, with elements $\boldsymbol{y}_i(t)$. Further exogenous explanators may exist at unit or dyadic level, $\mathbf{v}_{\mathrm{i}}(t)$ or $\mathbf{w}_{ij}(t)$—in our notation, the components of $\mathbf{X}$. Actors have opportunities to make changes in their network connections, switching on or off one tie or none, at fixed rate in continuous time, λ_i^{net}, according to an exponential hazard-rate model. The model may further parameterize λ, but the paper's implementation assumed the rate constant across all ij and t. Likewise, opportunities to switch the behavior on or off or do nothing occur at continuous-time rate λ_i^{beh}.[30]

[29] We follow Snijders et al. (1997) and Steglich et al. (2006, 2007) most specifically.

[30] Since observation occurs at discrete intervals, the degrees of freedom to vary these continuous-time rates

When the opportunity to change network ties arrives for some i, s/he chooses to change the status of any one of his/her *N-1* ties, turning it on or off, or leaving all ties unchanged. S/he makes this choice by comparing the values of some objective function of this form:

$$f_i^{net}(\mathbf{x},\mathbf{x}',\mathbf{z})+\boldsymbol{\varepsilon}_i^{net}(\mathbf{x},\mathbf{x}',\mathbf{z})=\sum\nolimits_h\left\{\boldsymbol{\beta}_h^{net}\times\mathbf{s}_h^{net}(\mathbf{i},\mathbf{x},\mathbf{x}',\mathbf{z})\right\}+\boldsymbol{\varepsilon}_i^{net}(\mathbf{x},\mathbf{x}',\mathbf{z}) \qquad (27),$$

where $\mathbf{x}'$ is an alternative network under consideration, which can differ from the existing network, $\mathbf{x}$, only by changing at most one element of (only) row i.[31] $s_h^{net}(\cdot)$ is some statistic, i.e., some function of the data, $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{z}$, that reflects the substantively/theoretically derived objectives of the actors with regard to the network, $\mathbf{x}$, and behaviors, $\mathbf{z}$. The $\boldsymbol{\beta}_h^{net}$ to be estimated are the relative weights of these objectives. Assuming the $\boldsymbol{\varepsilon}_i^{net}(\mathbf{x},\mathbf{x}',\mathbf{z})$ extreme-value distributed, independently across actors (see note 31) and over time, produces the multinomial-logit model of categorical choice. Similarly, when an opportunity to change behavior arrives, actor i compares the value of her/his objective function under each of the three possible actions: increment or decrement by one or leave unchanged. Formally, i compares $\mathbf{z}$ to $\mathbf{z}'$ given $\mathbf{x}$, and under analogous assumptions of *i.i.d.* extreme-value stochastic components, the multinomial logit again emerges.[32]

As in Leenders' approach, identification derives from debarring any literal simultaneity in outcomes or networks and assuming that temporal precedence implies causal precedence, and in particular conditioning on the first observation.[33] Given all this, estimation occurs by

---

render the assumption of exclusively single actors making single, unit-valued changes in their network connections or behavior essentially inconsequential. As greater frequency and/or magnitude of changes are observed, estimates of these occurrence rates at this unobserved instantaneous level simply rise to compensate. This does not, however relax the strong assumption of conditional independence of these actors' choices.

[31] The Steglich et al. (2007:21) exposition actually omits the stochastic component from the right-hand side of (27), and seems to carry this omission into the simulation-model implementation and the associated "method of moments" estimation. We suspect this is highly consequential because it suppresses the dependence across units or dyads of their multinomial choices (see note 33) regarding which if any $\mathbf{x}_{ij}$ to switch on or off.

[32] Given the binary behavior and the model set-up, we see only two possible choices: change the behavior's on or off status or leave it unchanged. In this case, the multinomial logit seems to reduce to the simple logit.

[33] Some identification problems persist with the current implementations, notwithstanding these strong assumptions. For one, assuming independent multinomial decisions for the endogenous behaviors and network ties and of opportunities for action effectively undoes some of the allowance for dependence in those choices, although it yields the great advantage of seeming to allow estimating standard multinomial logit (and exponential hazard-rate) models for those components of the system. That evasion aside, though, another issue is that included among the unit or dyadic explanators are various measures of network structure or units' places therein. These are functions of the ties between actors (and possibly also their behaviors), i.e., of the outcomes of the multinomial choices of the actors regarding the connections. In latent-variable models like the multinomial logit, however, one cannot include the actual outcomes on the right-hand side, however lagged or transformed by some network-structure measurement-function. Only the latent variable or the estimated probabilities can enter those functions. (The problem is that the actual outcomes indirectly enter

simulating the sequences of policies $\mathbf{z}$ and of networks $\mathbf{x}$ and searching over possible values of the model parameters, $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$, to minimize some distance function from the observed sequences of $\mathbf{x}$ and $\mathbf{z}$ to the simulated sequences. Snijders et al. (1997, 2007) label this as an application of 'the method of moments" and an example of a 'third-generation problem' in applied statistics (citing Gouriéroux & Monfort 1996 on the latter); one could also think of it as a calibration exercise. Standard errors could derive from jackknife or bootstrapped resampling (Snijders & Borgatti 1999) if explicit likelihoods or sufficient-statistic moment-equations are unavailable for standard analytic formulae.[34]

For C&IPE, some features of extant network-coevolution approaches, for all the valuable advances they offer, are not ideal as currently implemented. First, relative connectivity between units and many behaviors or attributes of interest as dependent variables in C&IPE are less likely to be binary or ordinal as current coevolution models require.[35] In the canonical globalization-and-tax-competition context, for instance, the outcomes of interest are tax rates, and many sources of interdependence will derive from the strength of economic competition. Second, in C&IPE contexts, strengths of relative connectivity are often unobserved, or even unobservable, directly. Continuing the example, we can observe only some covariates, like geographic contiguity and proximity or trade and capital flows, theorized to relate to the unobserved strength of economic competition. Thus, we would have no data for the left-hand side of the selection models in extant network-coevolution approaches. We could estimate the network and its determinants only by estimating their impact on actors' behavior given some spatial-econometric model of how the network matters for that behavior and how some observed covariates relate to network

the right-hand side to predict their own probabilities: see Heckman 1978.) The presence of a stochastic component exhibiting dependence across units, moreover, would render the multinomial logits N-dimensional optimization exercises rather than the standard unidimensional. We, however, have no further progress on those problems to offer here, beyond some conjectures we make in the conclusion, nor do we know of any other scholarship that has made greater progress on these issues in this behavior and network coevolution context. (Spatial econometricians have made considerable progress on this multidimensional optimization issue of interdependent latent-variable models, but entirely outside the co-evolution context to our knowledge.)

[34] The work we read indicated that these explicit likelihoods or proofs of the moment-equations sufficiency were not known, but, at least as of SIENA 3.17a (8 April 2008), estimated variance-covariance of the estimated parameters derive from the appropriate analytic calculations for moments or likelihood estimation.

[35] We suspect that Snijders et al.'s SIENA actually requires only discrete, not necessarily ordinal, behaviors. The limit seems the sensibility of conceiving an option to increment, leave unchanged, or decrement the behavior by one. If so, rounding or rescaling continuous behaviors to render them discrete should suffice. Unbounded behaviors would actually simplify by removing need to alter actors' choice problem at the bounds.

ties. Third, temporal precedence often will not suffice to assure causal precedence, as these models assume, for the many possible reasons reviewed above in the context of the time-lagged spatial-lag model. For one, interdependence often operates literally simultaneously in C&IPE. Most political-economy relations are strategic and, in strategic interactions, the effect of *alters* on *ego* is instantaneous or based in expectations. The interdependence of tax policies across units, for instance, arises from policymakers' simultaneous strategic play of a game in which the optimal policies of each actor depend on the current or expected-future policies of others. For another, *simultaneous* means within an observational period, and many C&IPE contexts have high frequency behavior and/or network changes relative to much lower observation periodicity. Furthermore, time lagging will suffice to eliminate simultaneity only if and insofar as these and other conditions discussed above apply. Finally, conditioning on the first observation loses least information and suffers least small-sample bias with long T, which does not frequently obtain in C&IPE.[36]

As Leenders (1997:165) underscores, most research on network/spatial dependence either studies the formation of networks (*selection*), taking actors' attributes and behaviors as fixed and given, or the effects on behaviors of networks/interdependence (*contagion*), taking the pattern of connectivity as fixed and given. Spatial econometricians have worked primarily in the latter mode, whereas network analysts have worked mostly in the former, although both are eager to combine the two. Other differences in tendency appear to us. Spatial econometricians tend primarily to conceive *network effects* as the effects of alters' actions on ego's via their connections, whereas network analysts tend to stress the effects of network structure and ego's position in it on actions, but this difference in core question—what explains networks vs. how interdependence affects outcomes—seems to us the most crucial. Among network analysts, Snijders and colleagues' coevolution model represents the greatest advances, to our knowledge, toward this combining of *contagion* and *selection*.

We approach coevolution from a spatial-econometric vantage and so start with the spatiotemporal-lag model, (5), and expand its specification to allow estimation of $\mathbf{W}$, the matrix of relative connectivity, modeling the w_{ij} as a parameterized function of covariates observable at unit, dyadic, or exogenous-external level. This model of the w_{ij} corresponds to the model of *selection* from the network-analytic view. E.g., the sociologists' *homophily* (like seeking/mimicking like), if it stems from fixed or exogenous characteristics of ego and alter,

[36] These obviously are general concerns, not necessarily restricted to C&IPE.

parallels a model from the spatial-econometric perspective of w_{ij} as a function of $\mathbf{x}_i$ and $\mathbf{x}_j$. If we consider some function of the vector of behaviors of interest, $\mathbf{y}$, among these explanators of $\mathbf{W}$, this parallels a stronger form of *selection*, raising higher statistical hurdles, in which network ties and actors' behaviors are jointly endogenous. Endogenous *homophily* would have w_{ij} some inverse-distance function of y_i and y_j, for example. Thus, the spatiotemporal-lag model with estimated, endogenous spatial-weights integrates contagion and selection in the spatial-econometric analogue to the network co-evolution model.

Consider $m\text{-}STAR$, a spatiotemporal-lag model with multiple spatial-weights matrices:

$$\begin{aligned}\mathbf{y} &= \rho_1\mathbf{W}_1\mathbf{y} + \rho_2\mathbf{W}_2\mathbf{y} + \ldots + \rho_R\mathbf{W}_R\mathbf{y} + \phi\mathbf{My} + \mathbf{X\beta} + \boldsymbol{\varepsilon} \\ &= \mathbf{Wy} + \phi\mathbf{My} + \mathbf{X\beta} + \boldsymbol{\varepsilon}, \text{ where } \mathbf{W} \equiv \sum_{r=1}^{R}\rho_r\mathbf{W}_r\end{aligned} \quad (28).$$

Notice that we can also write (28) in scalar notation as:

$$\begin{aligned}y_i &= \rho_1\sum_j w_{ij}^1 y_j + \rho_2\sum_j w_{ij}^2 y_j + \ldots + \rho_R\sum_j w_{ij}^R y_j + \phi y_{i,t-1} + \sum_k x_k^i\beta_k + \varepsilon_i \\ &= \sum_j\left(\rho_1 w_{ij}^1 + \rho_2 w_{ij}^2 + \ldots + \rho_R w_{ij}^R\right)y_j + \sum_k x_k^i\beta_k + \varepsilon_i \\ &= \sum_j\left\{\left(\sum_r \rho_R w_{ij}^R\right)y_j\right\} + \sum_k x_k^i\beta_k + \varepsilon_i\end{aligned} \quad (29).$$

As the middle line of (29) perhaps best clarifies, the term in parentheses is a parameterized (linear-additive) model of weights on $y_{j\neq i}$ in affecting y_i. The w_{ij}^r are the covariates offered to explain the pattern of interdependence, and ρ_r their coefficients to estimate. Thus, we can conceive the $m\text{-}STAR$ model as a $STAR$ model with the estimated $\mathbf{W}$, i.e., the estimated network, being a weighted sum of observed explanators of connectivity, $\hat{\mathbf{W}} = \sum_r \hat{\rho}_r\mathbf{W}_r$. If, furthermore, any $\mathbf{W}_r$ has functions of $\mathbf{y}$ as elements, then $\mathbf{W}$ and $\mathbf{y}$ are jointly endogenous, and (29) is a network-behavior coevolution model.

Without considerable further complication, the sorts of models of $\mathbf{W}$, i.e., of networks, expressible in this form are limited to those with continuous w_{ij} strengths of ties. If we expected truly binary connectivity, one would need to transform the term in parentheses to binary outcomes, say, by applying the log-odds function and a decision rule to convert probabilities to one. (This is not so great a limitation if one believes, as we tend to do, that connectivity is a degree, measured at best with error.) Other non linear-additive models of w_{ij} would also entail complications but seem manageable. Then, too, the costs in estimation complexity of enriching the model of connectivity by adding covariates is high compared to adding unit, dyad, or exogenous-external covariate $\mathbf{x}$ in $\mathbf{X\beta}$ (but perhaps not compared to extant network-coevolution models). The approach has some major advantages too though, notably among them that fully developed likelihoods for the m-STAR model exist, at least in the exogenous-$\mathbf{W}_r$ case, both the simpler likelihood conditional on the first observation and the unconditional one better-suited for *instantaneous* interdependence or small T. Thus, we can apply all the apparatus for estimation, all the analytically or simulation-derived intuitions about biases, efficiency, and sensitivity, and all the procedures for calculating, interpreting, and presenting spatiotemporally dynamic effects for the spatial-econometric models discussed above. On the other side, we can interpret and present the estimated network, $\hat{\mathbf{W}}$, conversely, with all standard network-analytic tools.

The conditional likelihood for *m-STAR* extends that of (11) for *STAR* intuitively:

$$\ln L(\boldsymbol{\rho},\boldsymbol{\beta},\phi,\sigma;\mathbf{y},\mathbf{X}) = \ln(2\pi\sigma^2)^{-NT/2} + \ln\left|\mathbf{A}\right| - \frac{1}{2\sigma^2}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}, \qquad (30).$$
$$\text{where } \mathbf{A} = \mathbf{I}_{NT} - \mathbf{W} \text{ and } \boldsymbol{\varepsilon} = \mathbf{Ay} - \phi\mathbf{My} - \mathbf{X\beta}$$

Written for (N×1) vectors $\mathbf{y}$, the likelihood is conveniently separable as follows, highlighting the conditionality on the first observation (which is not apparent in (30)):

$$\text{Log } f_{\mathbf{y}_t,\mathbf{y}_{t-1},\ldots,\mathbf{y}_2|\mathbf{y}_1} = -\frac{1}{2}N\left(T-1\right)\log\left(2\pi\sigma^2\right) + \left(T-1\right)\sum_{r=1}^{R}\log\left|\mathbf{I}-\rho_r\mathbf{W}_r\right| - \frac{1}{2\sigma^2}\sum_{t=2}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t \qquad (31).$$
$$\text{where } \boldsymbol{\varepsilon}_t = \mathbf{y}_t - \sum_{r=1}^{R}\rho_r\mathbf{W}_{rN}\mathbf{y}_t - \phi\mathbf{I}_N\mathbf{y}_{t-1} - \mathbf{X}_t\boldsymbol{\beta}$$

The unconditional (exact) likelihood extends the more complex (13) analogously. Luckily, T is large enough in our application that the more compact conditional likelihood is adequate. In either case, the estimated variance of $\hat{\mathbf{W}}_{ij}$ is:

$$\hat{\mathbf{W}} = \sum_r \hat{\rho}_r\mathbf{W}_r \Rightarrow \widehat{\text{Var}(\hat{\mathbf{W}}^{(i,j)})} = \left[\mathbf{W}_1^{(i,j)}\ \ \mathbf{W}_2^{(i,j)}\ldots\mathbf{W}_R^{(i,j)}\right]\hat{\boldsymbol{\Omega}}_{\hat{\boldsymbol{\rho}}}\left[\mathbf{W}_1^{(i,j)}\ \ \mathbf{W}_2^{(i,j)}\ldots\mathbf{W}_R^{(i,j)}\right]' \qquad (32),$$

where $\hat{\boldsymbol{\Omega}}_{\hat{\boldsymbol{\rho}}}$ is from minus the inverse of the Hessian of the likelihood in the usual fashion.

Coevolution models, i.e., models where (network) connectivity, $\mathbf{W}$, is some function of (behavior) $\mathbf{y}$, present larger challenges. Our simple stratagem for a first cut is to apply *the poor man's exogeneity*: we lag temporally the $\mathbf{y}$ in this function explaining $\mathbf{W}$ and assume the conditions required for that identification approach to hold sufficiently. As noted above, this does not address the problem of true or effective simultaneity, which seems likely in C&IPE contexts at least. Therefore, we are also currently exploring a two-step estimation procedure. First, apply spatial-GMM to obtain consistent estimates of the endogenous w_{ij}

and their estimated variance-covariance by instrumentation. Then, draw the $\hat{\mathbf{W}}$ to insert in the likelihood, conditional (30) or the unconditional extension of (13), from this estimated multivariate distribution. Maximize these q likelihoods, each time with new draws from that first-stage S-GMM instrumented $\hat{\mathbf{W}}$. The average of the q second-stage S-ML estimates provide point estimates of parameters, and the estimated variance-covariance of those parameter-estimates is the average of the estimated variance-covariance matrices from each iteration plus $(1+q)$ times the sample variance-covariance in the point estimates across iterations (as, e.g., in King et al. 2001 multiple imputation). This estimator should inherent nice properties from S-GMM and S-ML as far as we can intuit, although we have neither analytic nor simulation demonstration of properties yet.[37] Assessment of the estimator and direct comparison to network-coevolution approaches would then be essential next steps.

IV. Monte Carlo Simulation of S-ML vs. S-OLS vs. OLS Estimation of m-STAR Models

Before illustrating the estimation, testing, and interpretation of our m-STAR model of network-behavior coevolution, we will demonstrate that, in fact, the S-ML estimators just described are needed and outperform simpler least-squares estimators. Analytically, the omitted-variable biases of the blind OLS estimator remain as before: $\mathbf{F\beta}$. The simultaneity asymptotic biases (inconsistencies) of the naïve S-OLS estimator, which simply inserts the multiple spatial lags into a least-squares regression, are also analogous to (8) as follows:

$$\text{Let } \mathbf{y} = \mathbf{Z\rho} + \boldsymbol{\varepsilon} \;\; ; \;\; \text{where } \mathbf{Z} = \begin{bmatrix} \mathbf{W}_1\mathbf{y} & \mathbf{W}_2\mathbf{y} \end{bmatrix} \text{ and } \boldsymbol{\rho} = \begin{bmatrix} \rho_1 & \rho_2 \end{bmatrix}'.$$

$$\text{Then: plim } \hat{\boldsymbol{\rho}} = \boldsymbol{\rho} + \text{plim}\left[\left(\frac{\mathbf{Z}'\mathbf{Z}}{n}\right)^{-1}\frac{\mathbf{Z}'\boldsymbol{\varepsilon}}{n}\right]\text{; that is:}$$

$$\text{plim } \hat{\boldsymbol{\rho}} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} + \begin{bmatrix} \dfrac{\text{var}(\mathbf{W}_2\mathbf{y})}{\text{var}(\mathbf{W}_2\mathbf{y})\text{var}(\mathbf{W}_1\mathbf{y}) - [\text{cov}(\mathbf{W}_1\mathbf{y},\mathbf{W}_2\mathbf{y})]^2} & \dfrac{-\text{cov}(\mathbf{W}_1\mathbf{y},\mathbf{W}_2\mathbf{y})}{\text{var}(\mathbf{W}_2\mathbf{y})\text{var}(\mathbf{W}_1\mathbf{y}) - [\text{cov}(\mathbf{W}_1\mathbf{y},\mathbf{W}_2\mathbf{y})]^2} \\ \dfrac{-\text{cov}(\mathbf{W}_1\mathbf{y},\mathbf{W}_2\mathbf{y})}{\text{var}(\mathbf{W}_2\mathbf{y})\text{var}(\mathbf{W}_1\mathbf{y}) - [\text{cov}(\mathbf{W}_1\mathbf{y},\mathbf{W}_2\mathbf{y})]^2} & \dfrac{\text{var}(\mathbf{W}_1\mathbf{y})}{\text{var}(\mathbf{W}_2\mathbf{y})\text{var}(\mathbf{W}_1\mathbf{y}) - [\text{cov}(\mathbf{W}_1\mathbf{y},\mathbf{W}_2\mathbf{y})]^2} \end{bmatrix} \begin{bmatrix} \text{cov}(\boldsymbol{\varepsilon},\mathbf{W}_1\mathbf{y}) \\ \text{cov}(\boldsymbol{\varepsilon},\mathbf{W}_2\mathbf{y}) \end{bmatrix}; \quad (33).$$

$$\text{that is: plim } \hat{\boldsymbol{\rho}} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} + \begin{bmatrix} \dfrac{\Psi_{22}\Gamma_{11} - \Psi_{12}\Gamma_{21}}{|\boldsymbol{\Psi}|} \\ \dfrac{\Psi_{11}\Gamma_{21} - \Psi_{12}\Gamma_{11}}{|\boldsymbol{\Psi}|} \end{bmatrix}, \text{ where } \boldsymbol{\Psi} = \text{plim}\left(\frac{\mathbf{Z}'\mathbf{Z}}{n}\right) \text{ and } \boldsymbol{\Gamma} = \text{plim}\left(\frac{\mathbf{Z}'\boldsymbol{\varepsilon}}{n}\right)$$

The intuitions remain as before: simultaneity biases generally increase in ρ, and they induce biases in generally opposite directions for other covariates' coefficient-estimates.

To demonstrate that an estimator is inconsistent, however, does not demonstrate that

[37] Neither do we as-yet have adequately functioning proof of concept, in fact.

these inconsistencies are practically large or that they outweigh other potential deficiencies of consistent estimators. Accordingly, we conduct some simple Monte Carlo simulations to explore the small-sample performance of these estimators, specifically the magnitudes of their biases, inefficiency, and standard-error inaccuracy.

Table 1. Estimator Comparison for m-STAR Model: S-OLS vs. S-ML

Estimator	**Result**	$\boldsymbol{\beta_0}$**=1**	$\boldsymbol{\beta_1}$**=1**	$\boldsymbol{\rho_1}$**=.3**	$\boldsymbol{\rho_2}$**=.3**
S-OLS	*Average Estimate*	.38 / .24	.96 / .97	.47 / .47	.27 / .29
	Standard Deviation	.33 / .24	.07 / .05	.19 / .15	.21 / .16
	Root Mean-Squared-Error	.71 / .80	.08 / .06	.25 / .22	.21 / .16
	Average Std-Err Estimate	.37 / .28	.06 / .05	.15 / .11	.17 / .13
	Overconfidence	.92 / .87	1.06 / 1.02	1.29 / 1.35	1.21 / 1.27
S-ML	*Average Estimate*	1.09 / 1.05	1.00 / 1.00	.31 / .31	.27 / .28
	Standard Deviation	.33 / .24	.07 / .05	.12 / .09	.14 / .11
	Root Mean-Squared-Error	.34 / .24	.07 / .05	.12 / .09	.14 / .11
	Average Std-Err Estimate	.31 / .23	.06 / .05	.12 / .09	.14 / .11
	Overconfidence	1.05 / 1.05	1.03 / 1.00	.98 / 1.01	.98 / 1.01

Monte Carlo (1000 Trials) Results for $\mathbf{y}=\rho_1\mathbf{W}_1\mathbf{y}+\rho_2\mathbf{W}_2\mathbf{y}+\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$, with $\mathbf{W}_1$=rook adjacency, $\mathbf{W}_2$=queen adjacency (row normalized); $\beta_0=\beta_1=1$, $\rho_0=\rho_1=.3$; and N=225/450.

The results in Table 1 are easily interpreted using the analytical results in (33). The covariance of the *queen* spatial-lag (all eight adjacent squares on a grid) and $\boldsymbol{\varepsilon}$ is much less than that of the *rook* lag (only the four horizontally and vertically adjacent) and $\boldsymbol{\varepsilon}$. With row-standardization and eight connections the strength of the interdependence/endogeneity is diluted in the former case. Consequently, the coefficient on rook-lag is overestimated, and the coefficients on *queen*-lag $\mathbf{W}_2\mathbf{y}$ and on x_0 (which is especially correlated with $\mathbf{W}_1\mathbf{y}$)[38] are (badly) underestimated. The S-ML estimator also dominates impressively in efficiency and standard-error accuracy, especially for those two estimates $\hat{\rho}_1$ and $\hat{\beta}_0$.

Next, Table 2 similarly evaluates the blind OLS, naïve S-OLS, and S-ML estimators for our m-STAR coevolution model. The S-ML estimator again outperforms the inconsistent OLS alternatives. Notice that, with $\mathbf{x}$ drawn independently, appreciable correlation of the regressors with the spatial lags concentrates in the time-lag; thus, omitted-variable biases of blind OLS are not severe for β and concentrate at a noticeable 20% in ϕ. Notice also that with endogenous $\mathbf{Ly}$ being time-lagged in the estimator and in truth, and $\mathbf{L}$ being $|y_i-y_j|$ and so not terribly (linear) correlated with y_i, the simultaneity biases of S-OLS concentrate in ρ at a sizable +33% but, $\mathbf{x}$ being drawn independently, induce little bias in $\hat{\beta}$. The efficiency

[38] The simulations drew $\boldsymbol{\varepsilon}$ from N(1,1), making the nonzero aspect of $\mathbf{Z'Wy}$ concentrate in the constant, $\mathbf{x}_0$.

(RMSE) and standard-error accuracy gains are more uniform and obvious.

Table 2. Estimator Comparison for m-STAR Coevolution Model: OLS v. S-OLS v. S-ML

ESTIMATOR	RESULT	ϕ=.3	β=1	ρ=.3	γ=.3
OLS	*Average Estimate*	.36	1.04	—	—
	Standard Deviation	.07	.05	—	—
	Root Mean-Squared-Error	.09	.06	—	—
	Average Std-Err Estimate	.03	.05	—	—
	Overconfidence	2.07	.97	—	—
S-OLS	*Average Estimate*	.28	.99	.41	.25
	Standard Deviation	.03	.05	.06	.08
	Root Mean-Squared-Error	.04	.05	.129	.10
	Average Std-Err Estimate	.03	.04	.05	.08
	Overconfidence	1.02	1.05	1.19	1.01
S-ML	*Average Estimate*	.29	1.00	.31	.27
	Standard Deviation	.03	.05	.05	.08
	Root Mean-Squared-Error	.03	.05	.05	.09
	Average Std-Err Estimate	.03	.04	.04	.07
	Overconfidence	1.02	1.04	1.11	1.10

Monte Carlo (1000 Trials) Results for $\mathbf{y}_t=\rho\mathbf{W}\mathbf{y}_t+\gamma\mathbf{L}\mathbf{y}_{t-1}+\phi\mathbf{y}_{t-1}+\beta\mathbf{x}_t+\boldsymbol{\varepsilon}_t$, with $\mathbf{W}$=48 contiguous US-state adjacency pattern (row-stdzd); ρ=.3, γ=.3, ϕ=.3, β=1; and N=48, T=10.

In sum, even in simulations rather favorably designed for the blind or naïve estimators, the S-ML estimator is clearly dominant for all estimates and estimate-properties.

V. Empirical Illustration

To illustrate application of the S-ML estimated m-STAR approach to endogenous network-behavior coevolution (with identification from temporal lagging assumed), we extend our previous ALM-policy analysis (Franzese & Hays 2006c). One extension is of the sample to include observations on both *EU* and non-*EU* countries.[39] This allows distinction of co-membership interdependence among *EU* member states from global interdependence.

The OECD ALM-program dataset gives expenditures by five categories: labor-market training, public employment-services and administration, subsidized employment, youth measures, and disability measures. Figure 1 plots the temporal variation in *OECD* average spending by type. Subsidized employment and labor-market training are the two largest components over the entire sample period, accounting for 26.9% and 26.7% the total.

[39] Annual 1980-2003 data for 21 OECD countries, 14 being EU members: Australia, Austria, Belgium, Canada, Denmark, Germany, Greece, Finland, France, Ireland, Italy, Japan, New Zealand, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the U.K., and the U.S.

Figure 1: Disaggregated Active Labor Market Expenditures in the *OECD*, 1980-2003

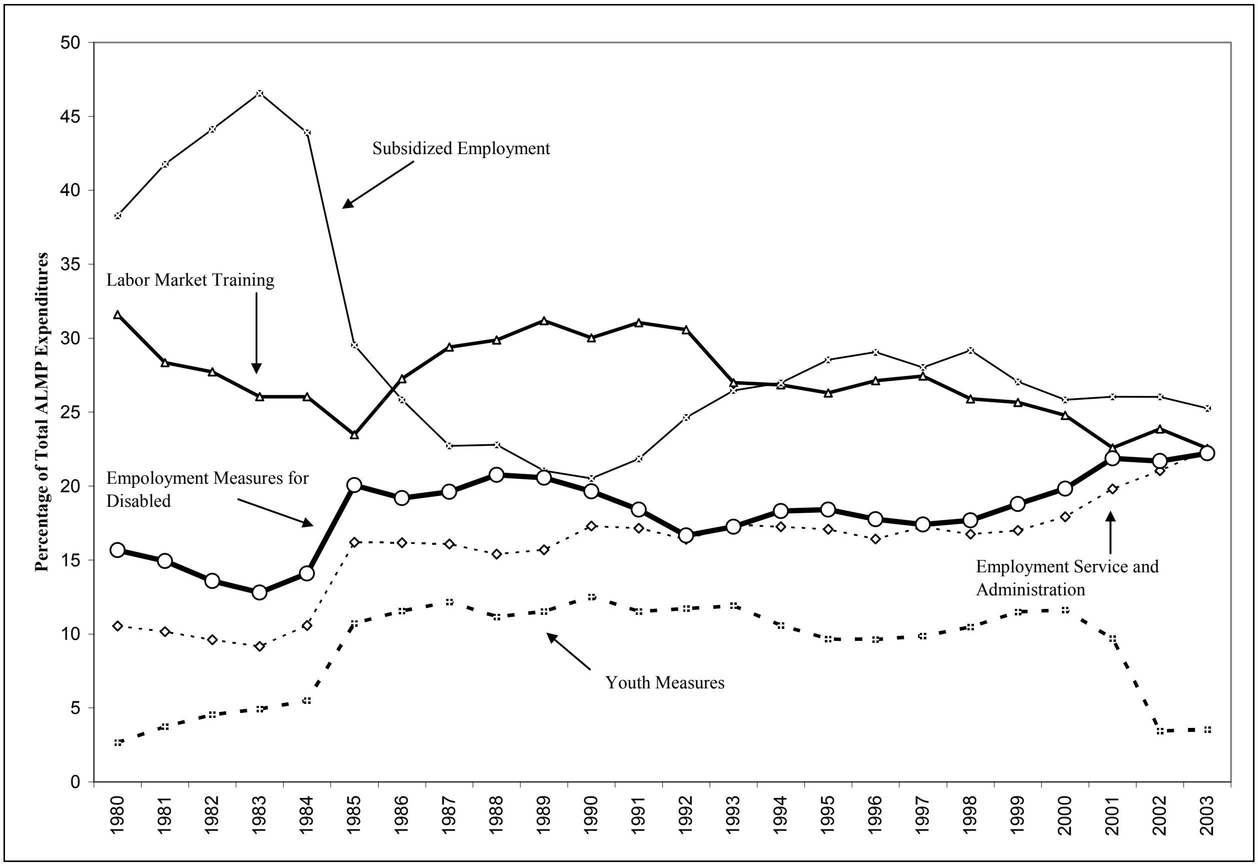


Table 3 gives the programmatic breakdown in ALM expenditures by country, revealing great variation across the sample. The big spenders per capita were Sweden and Denmark (\$360.88 and \$287.20 (2000, PPP\$)); the U.S. and Greece spent least (\$43.72 and \$34.97). The table also reveals some spatial clustering on geographic, cultural, institutional-structure dimensions: e.g., all four Scandinavian countries spent much more than the OECD average; Portugal and Spain averaged within \$1 per capita of each other over these 23 years; and Australia and New Zealand, Canada and the U.S. spent well below the OECD average. What explains these patterns: strategic policy interdependence, similar exogenous-external conditions, correlated domestic factors, or some selection process among countries grouped on these dimensions? Which dimensions?

Table 3. Disaggregated Active Labor Market Expenditures per Capita (2000 PPP$)											
	AUS	*AUT*	*BEL*	*CAN*	*DEN*	*FIN*	*FRA*	*DEU*	*GRE*	*IRE*	*ITA*
Employment services&admin	28.85 (37.54)	33.06 (29.68)	41.35 (15.69)	38.44 (39.1)	20.48 (7.13)	22.47 (10.46)	25.98 (14.39)	46.59 (19.12)	10.77 (30.8)	32.36 (15.19)	0 (0)
Labor-market training	10.05 (13.08)	41.58 (37.32)	43.99 (16.69)	46.18 (46.98)	117.32 (40.85)	67 (31.19)	49.74 (27.56)	71.61 (29.38)	7.84 (22.41)	43.26 (20.3)	7.14 (8.02)
Youth measures	8.36 (10.88)	4.63 (4.16)	3.41 (1.29)	3.92 (3.99)	24.57 (8.55)	17.65 (8.22)	42.47 (23.53)	12.83 (5.26)	6.8 (19.45)	37.11 (17.42)	31.47 (35.38)
Subsidized employment	21.51 (27.98)	18.62 (16.71)	146.9 (55.73)	7.07 (7.19)	69.26 (24.12)	87.79 (40.87)	49.75 (27.56)	65.7 (26.96)	8.28 (23.67)	87.35 (41)	24.68 (27.74)
Measures for disabled	8.08 (10.52)	13.51 (12.13)	27.94 (10.6)	2.7 (2.75)	55.58 (19.35)	19.12 (8.9)	12.56 (6.96)	46.95 (19.27)	1.28 (3.67)	13 (6.1)	0 (0)
Total ALMP	76.87	111.41	263.61	98.3	287.2	214.81	180.51	243.72	34.97	213.08	88.97
	JPN	*NTH*	*NWZ*	*NOR*	*PRT*	*ESP*	*SWE*	*CHE*	*GBR*	*USA*	*OECD*
Employment services&admin	47.65 (68.44)	24.48 (11.58)	17.46 (15.85)	33.36 (15.91)	14.93 (17.04)	12.97 (14.69)	44.98 (12.46)	20.81 (22.57)	36.4 (37.71)	12.1 (27.67)	26.76 (17.34)
Labor-market training	7.48 (10.75)	48.72 (23.04)	39.62 (35.96)	36.28 (17.31)	26.47 (30.22)	19.23 (21.78)	101.86 (28.22)	17.01 (18.45)	15.5 (16.06)	13.48 (30.82)	41.12 (26.65)
Youth measures	0.24 (0.34)	11.01 (5.21)	10.44 (9.48)	11.53 (5.5)	26.48 (30.23)	7.73 (8.76)	14.32 (3.97)	0.55 (0.59)	28.34 (29.36)	6.68 (15.29)	14.84 (9.61)
Subsidized employment	12.73 (18.28)	16.69 (7.89)	34.73 (31.53)	23.87 (11.39)	14.77 (16.86)	45.7 (51.77)	91.75 (25.42)	20.05 (21.74)	11.68 (12.1)	2.24 (5.13)	41.44 (26.86)
Measures for disabled	1.53 (2.19)	110.57 (52.28)	7.91 (7.18)	104.59 (49.89)	4.95 (5.65)	2.64 (2.99)	107.98 (29.92)	33.79 (36.65)	4.61 (4.77)	9.22 (21.09)	29.29 (18.98)
Total ALMP	69.62	211.47	110.15	209.63	87.59	88.28	360.88	92.22	96.52	43.72	154.3

Note: Parentheses contain spending as a percentage of total spending on active labor market programs.

To answer these questions, we estimate an m-STAR model with coevolutionary dynamics.[40] The model, in matrix notation, is

$$\mathbf{y} = \left[\sum_{r=1}^{R} \rho_r \mathbf{W}_r\right]\mathbf{y} + \phi\mathbf{My} + \gamma\mathbf{Ly} + \mathbf{X\beta} + \boldsymbol{\varepsilon} \quad (34),$$

where $\mathbf{y}$, ALM expenditures, is an $NT\times 1$ vector of cross sections stacked by periods as described previously. ρ_r is the r^{th} spatial autoregressive coefficient, and $\mathbf{W}_r$ is an $NT\times NT$ block-diagonal spatial-weighting matrix. Each $\mathbf{W}_r$ contains a unique set of elements w_{ij}^r that reflect a particular type of interdependence (e.g., geographic proximity, *EU* co-membership, and economic interdependence). In the other new term, γ, is the coevolutionary-dependence parameter, and $\mathbf{L}$ is an $NT\times NT$ "policy-distance" matrix with $|y_{i,t-1}-y_{j,t-1}|$ in cells (it,jt). The addition of $\gamma\mathbf{Ly}$, therefore, reflects a substantive proposition that states with more-similar ALM policies (spending-levels, to be precise) affect each other's ALM policies more ($\gamma>0$), as in the network analyst's *homophily*, or less ($\gamma<0$) than do states with less-similar

[40] Case et al. (1993), Brueckner & Saavedra (2001), Fredriksson & Millimet (2002), Redoano (2003), Allers & Elhorst (2005) among others have used spatial-lag models to test similar strategic policy-interdependence hypotheses, but none use multiple spatial lags or consider coevolution as alternative connectivity dimensions.

policies. In spatial-econometric terms, it is the endogenous determinant of the strength of interdependence. The rest is as before: $\mathbf{Wy}$ reflects the exogenous interdependence of units' policies; $\mathbf{My}$ is the first-order temporal lag, with ϕ its coefficient; $\mathbf{X\beta}$ are the exogenous non-spatial components; and $\boldsymbol{\varepsilon}$ are assumed-$i.i.d.$ disturbances.

The presence in (34) of $\mathbf{L}$, which contains lagged y's, renders the system of N equations nonlinear in the endogenous variable. This complicates calculation of spatiotemporal dynamics and prevents linear multipliers or analytical solution for steady states. The spatiotemporal coevolutionary responses to changes in $\mathbf{X}$ or $\boldsymbol{\delta}$ must be calculated recursively. To start, rewrite (34) as t cross-sections:

$$\mathbf{y}_t = \mathbf{W}_t\mathbf{y}_t + \phi\mathbf{y}_{t-1} + \gamma\left[abs\left(\mathbf{\Pi}\left[\mathbf{y}_{t-1}\otimes\mathbf{I}_N\right]\right)\right]\mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \qquad (35).$$

$\mathbf{y}_t$, $\mathbf{W}_t$, and $\mathbf{X}_t$ are N×1, N×N, and N×k matrices, and $\mathbf{\Pi}$ is an N×N^2 matrix produced by horizontally concatenating N separate N×N matrices. The i^{th} N×N matrix in $\mathbf{\Pi}$ has -1's on its diagonal and 1's for each element of the i^{th} column except for (i,i) which is 0, as are all remaining elements. If N=3, for example:

$$\mathbf{\Pi} = \begin{bmatrix} 0 & 0 & 0 & -1 & 1 & 0 & -1 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & -1 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix} \qquad (36).$$

The function abs is element-by-element absolute value; its argument gives the vector of differences $y_{i,t-1}$-$y_{j,t-1}$, reflecting *homophily.* In reduced form, (35) is

$$\mathbf{y}_t = \left[\mathbf{I}_n - \mathbf{W}_t + \gamma\left[abs\left(\mathbf{\Pi}\left[\mathbf{y}_{t-1}\otimes\mathbf{I}_N\right]\right)\right]\right]^{-1}\left(\phi\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right) \qquad (37).$$

Our empirical analysis applying (37) focuses on aggregate ALM-program expenditures and the two largest components thereof, subsidized employment (SEMP) and labor-market training (LMT) spending. Our dependent variables are measured per capita (2000, PPP$), and the key right-hand-side variables, which allow us to evaluate the nature of the spatial interdependence among the countries in our sample, are the spatial lags of ALM spending.

Our spatial lags, $\mathbf{W}_r\mathbf{y}$, involve four different weights matrices (R = 4). $\mathbf{W}_1$ is a standardized *binary contiguity-weights matrix* which begins by coding w_{ij}=1 for countries i and j that share a border and w_{ij}=0 for countries that do not border, with exceptions France, Belgium, and Netherlands treated as contiguous with U.K., Denmark with Sweden, and New Zealand with Australia. $\mathbf{W}_2$ is an EU co-membership weights-matrix; i.e., w_{ij}=1 if both i and j are EU members and w_{ij}=0 otherwise. $\mathbf{W}_3$ has weights that reflect the trade (imports+exports) shares between sample countries.[41] $\mathbf{W}_4$ is the policy-distance coevolution matrix with $|y_{i,t-1}$-$y_{j,t-1}|$ in cells (it,jt) as described above. For estimation, we *row-standardize* (as is common in spatial econometrics) all $\mathbf{W}$ matrices, dividing each cell by its row's sum.[42]

We also control several domestic variables. We include macroeconomic performance with real GDP-growth and unemployment rates. As economies grow wealthier, governments might provide more public goods and services, suggesting a positive coefficient estimate for GDP per capita: Wagner's Law. Alternatively, Baumol's Disease, which refers to an argued decreasing relative productivity in service sectors rendering finance of public services more difficult as economies develop and shift toward service-sector employment, would suggest a relation of wealth to ALM spending. Most likely, though, our GDP-*growth* measure will capture pseudo-automatic programmatic responses to macroeconomic cycles, suggesting a negative coefficient. Unemployment should receive a positive coefficient for the same reason.

Next, we control several structural features of a country's economy related to its labor markets and exposure to external shocks. The labor-market factors are union density and Iversen & Cusak's (2000) deindustrialization measure. Higher union density increases the influence of organized labor, so we expect it to associate closely with greater ALM spending. Iversen & Cusak (2000) argued that workers cross significant skill barriers when they shift from manufacturing and agriculture to services. Thus, we expect deindustrialization to spur LMT also. Many scholars argue that international economic exposure favors increased government spending, especially on programs that help workers adjust to external shocks (e.g., Ruggie 1982; Cameron 1978; Katzenstein 1985; Rodrik 1997; Hays et al. 2005). Others argue that increased international exposure produces competitive pressures that lead to smaller governments, but this mechanism is properly reflected in our model by the third spatial-lag (see Basinger & Hallerberg 2004, Franzese & Hays 2003, 2004b, 2005ab, 2006c, 2007ab, 2008c). We use trade openness as our measure of exposure.

We also include the working-age percentage of the population, the percent of cabinet seats held by left and Christian Democratic parties, and the percent of general-election

[41] More specifically, $w_{ij}(t)$ equals the sum of exports i to j and j to i and of imports i from j and j from i, divided by four times i's GDP. We use all four values because the data exhibit slight discrepancies between, e.g., i to j exports and j from i imports.

[42] Row normalization is not necessarily substantively neutral (see, e.g., Plümper & Neumayer 2008).

votes won by left-libertarian parties. Working-age voters are natural constituencies for ALM programs, whereas the benefits for retired voters are indirect at best, so political pressure for ALM policies should increase with working-age population-shares. Scholars have variously identified Social Democratic, Christian Democratic, and Left-Libertarian parties as key supporters of active social-policies, albeit of/to/for different precise natures, extents, or reasons (see, e.g., Garrett 1998; Swank 2002; and Kitschelt 1994). The simpler left-right ideological dimension may also relate to ALM programs.

Table 4: ALM-Spending Models — Estimation Results

DEP. VAR. →		**Total ALM**			**LMT**			**SEMP**	
INDEP. VAR. ↓	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
Constant	-195.89*	-150.369	51.487	-162.53***	-160.30**	-100.950	-122.27**	-129.70**	-65.945
	(103.645)	(114.470)	(125.395)	(59.486)	(64.481)	(69.774)	(57.794)	(61.645)	(71.708)
Temporal Lag	0.875***	0.836***	0.801***	0.872***	0.865***	0.758***	0.830***	0.848***	0.704***
	(0.026)	(0.032)	(0.0407)	(0.028)	(0.031)	(0.040)	(0.034)	(0.044)	(0.062)
Real GDP Growth Rate	1.365***	0.449	-1.100	0.269	0.237	-0.861	0.036	-0.056	-1.044*
	(0.411)	(0.607)	(1.144)	(0.201)	(0.258)	(0.630)	(0.188)	(0.231)	(0.629)
Stdzd Unemployment Rate	-0.070	0.552	0.125	0.361	0.844*	0.541	-0.169	-0.351	-0.544
	(0.791)	(0.835)	(0.876)	(0.445)	(0.456)	(0.509)	(0.438)	(0.448)	(0.490)
Union Density	0.888***	0.711**	0.527*	0.205	0.138	0.259	0.567***	0.645***	0.629***
	(0.302)	(0.318)	(0.321)	(0.167)	(0.176)	(0.176)	(0.169)	(0.175)	(0.184)
Deindustrialization	1.259	1.209	0.499	0.106	0.105	-0.183	1.426***	1.440***	1.0514**
	(0.785)	(0.776)	(0.773)	(0.449)	(0.447)	(0.435)	(0.438)	(0.438)	(0.440)
Trade Openness	-0.522***	-0.183	-0.192	-0.187*	-0.083	-0.009	0.018	0.039	0.104
	(0.169)	(0.196)	(0.213)	(0.097)	(0.114)	(0.118)	(0.104)	(0.113)	(0.133)
Working-Age Population	0.946	0.216	0.108	2.239***	2.113**	2.623***	-0.139	-0.075	0.189
	(1.498)	(1.603)	(1.629)	(0.863)	(0.909)	(0.940)	(0.829)	(0.860)	(0.934)
Left Cabinet-Seats	-0.024	0.001	0.001	0.035	0.044**	0.057***	-0.049**	-0.047**	-0.037*
	(0.039)	(0.039)	(0.038)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Christian Dem. Cab-Seats	-0.160*	-0.102	-0.085	-0.070	-0.074	-0.042	-0.032	-0.031	-0.037
	(0.095)	(0.096)	(0.093)	(0.054)	(0.053)	(0.052)	(0.053)	(0.053)	(0.052)
Left-Libertarian Vote	-0.285	-0.549	-0.728	-0.316	-0.535	-0.714**	-0.248	-0.208	-0.147
	(0.622)	(0.611)	(0.598)	(0.355)	(0.361)	(0.352)	(0.346)	(0.344)	(0.340)
SPATIAL WEIGHTS:									
Borders		-0.112***	-0.098***		-0.185***	-0.188***		0.058	0.054
		(0.040)	(0.040)		(0.057)	(0.060)		(0.054)	(0.056)
EU Membership		-0.071***	-0.060*		-0.079	-0.183**		-0.058	-0.097
		(0.034)	(0.036)		(0.061)	(0.073)		(0.062)	(0.065)
Trade Shares		0.386***	0.239**		0.316***	0.180		-0.017	-0.097
		(0.101)	(0.111)		(0.119)	(0.159)		(0.148)	(0.169)
Policy Distance		-0.053	-0.189***		-0.004	-0.298***		0.049	-0.205**
		(0.039)	(0.067)		(0.046)	(0.085)		(0.049)	(0.087)
TIME DUMMIES?	*No*	*No*	*Yes*	*No*	*No*	*Yes*	*No*	*No*	*Yes*
σ	21.221***	20.741***	19.564***	12.061***	11.828***	11.161***	11.794***	11.702***	11.154***
	(0.787)	(0.775)	(0.722)	(0.454)	(0.451)	(0.426)	(0.437)	(0.434)	(0.414)
Log-Likelihood	-1646.41	-1638.99	-1616.95	-1438.49	-1432.5	-1413.61	-1430.23	-1427.63	-1410.92

Note: All regressions include fixed country effects; models (3), (6) and (9) also included fixed year effects. All the spatial weights matrices are row-standardized. The parentheses contain standard errors.

*** Significant at the .01 level; ** Significant at the .05 level; * Significant at the .10 level.

Table 4 presents our results. We estimate three regressions for each dependent variable: the first includes country indicators but omits spatial lags, ignoring interdependence; the second includes the spatial lags and the country indicators; the third includes spatial lags and country and time indicators. The period dummies are a flexible way to model common OECD-wide trends and/or common (random) shocks in ALM expenditures. Recall that the most important issue methodologically in obtaining good estimates of the strength of interdependence, i.e., of $\boldsymbol{\rho}$, is to model well any alternative mechanisms by which the outcomes might correlate spatially, such as common exogenous shocks (e.g., global economic conditions) or correlated domestic factors. From that perspective, the country and year dummies serve as a powerfully conservative way to account for (near-)common outside shocks or spatially correlated (near-common) domestic factor. Failure to account for such alternatives will bias spatial-lag coefficient-estimates, usually positively.

The non-spatial and spatial model estimates suggest subtly differing explanations for the spatiotemporal patterns in total ALM expenditures. The non-spatial model points to domestic real GDP-per-capita growth, indicating strong procyclicality to total ALM though not LMT or SEMP components, and to labor-market structures, deindustrialization and especially union density. The spatial models suggest that the effects that the non-spatial models attribute to domestic growth and trade exposure seem instead to reflect spatial diffusion of responses global conditions. More interestingly perhaps, all three estimation techniques find sizable differences in sources of LMT versus SEMP spending. LMT seems closely related to workforce age-demographics and not very closely related to our labor-market structural or institutional measures. SEMP, to the contrary, counts strong labor and deindustrialization among its sources, and not age demography. The spatial models controlling for common shocks also show ALM policy, especially SEMP, more counter-cyclical to the domestic economy. Perhaps most interestingly, the spatial models suggest that, while total ALM spending is not particularly partisan, the composition is decidedly so, with LMT associated positively and SEMP negatively with left cabinets. Wald tests of the spatial-lag coefficients reveal strong evidence of interdependence in ALM policy, the *t*-tests on 13 of the 24 ρ_r estimates being significant at conventional levels and the six joint tests of the four ρ_r estimates *per* spatial model all overwhelmingly rejecting null hypotheses of zero coefficients, i.e., of the nonspatial model.[43] In particular, total ALM spending seems strongly

[43] Likelihood-ratio tests of the models and information criteria also strongly favor the spatial models.

spatially interdependent along all four dimensions of proximity, SEMP much less so, and LMT intermediately. Consequently, coefficient estimates in non-spatial models, especially of total ALM, will almost certainly suffer from omitted-variable bias.

We focus, therefore, on the spatial models of total ALM spending and, in particular, on the most-conservative time-dummies version. This model (column 3) finds few strong and significant effects of domestic conditions net of interdependence, common shocks, and fixed country-specific factors. Point estimates suggest positive ALM-spending response to unemployment, union density, deindustrialization, and working-age population, but only the response to union density is significant and sizable. They show negative responses to real-GDP growth, trade exposure, Christian-Democratic cabinet-seats, and Left-Libertarian vote-shares, but only the last and the countercyclical response to domestic growth are close to significant and sizable. No response at all to left cabinet-seats emerges, though we have already noted that this seems to mask a strongly left-partisan shifting from SEMP to LMT in ALM-spending composition. The estimated pattern of interdependence, i.e., the *implicit net network*, uncovers strongly negative interdependence among bordering countries and moderately negative interdependence among EU countries. The sign and relative strengths of interdependence by these patterns are consistent with our positive-externalities argument (Franzese & Hays 2006c). The negative $\hat{\rho}$ for the EU-membership spatial-lag also bolsters the case for those concerned that the EU is not adequately spurring employment-policy coordination. The positive coefficient for the trade-weights spatial-lag, meanwhile, supports arguments of globalization-induced competition. The coefficient(s) on the policy-distance lag(s) are negative and quite significant for total ALM (and for LMT and SEMP also). This indicates lesser dependency of domestic ALM policies on countries with more dissimilar ALM policies. I.e., policymakers follow more closely those more similar to them, as revealed by the similarity in the policies they choose: *homophily* in other (network-analytic) terms.

We are satisfied that ALM policy exhibits statistically significant interdependence, and that the patterns of interdependence relate to geographic contiguity and EU co-membership in ways that indicate policy free-riding, to trade relations in a way that implicates globalization and policy-competition, and to policy distance in a way that suggests homophily, but what do these results tell us of the net sign and substantive magnitude of this implicit network or of the effects of some countries' ALM policies on policymakers in other countries *via* this estimated implicit (net) network? What do they say about the

ALM-policy responses across these interdependent political-economies over time to counterfactual shocks in domestic and/or foreign conditions or policies? Answers and fuller interpretation of the coevolutionary spatiotemporal effects and dynamics that these m-STAR model-estimates imply requires calculation of the spatial multiplier in (37).

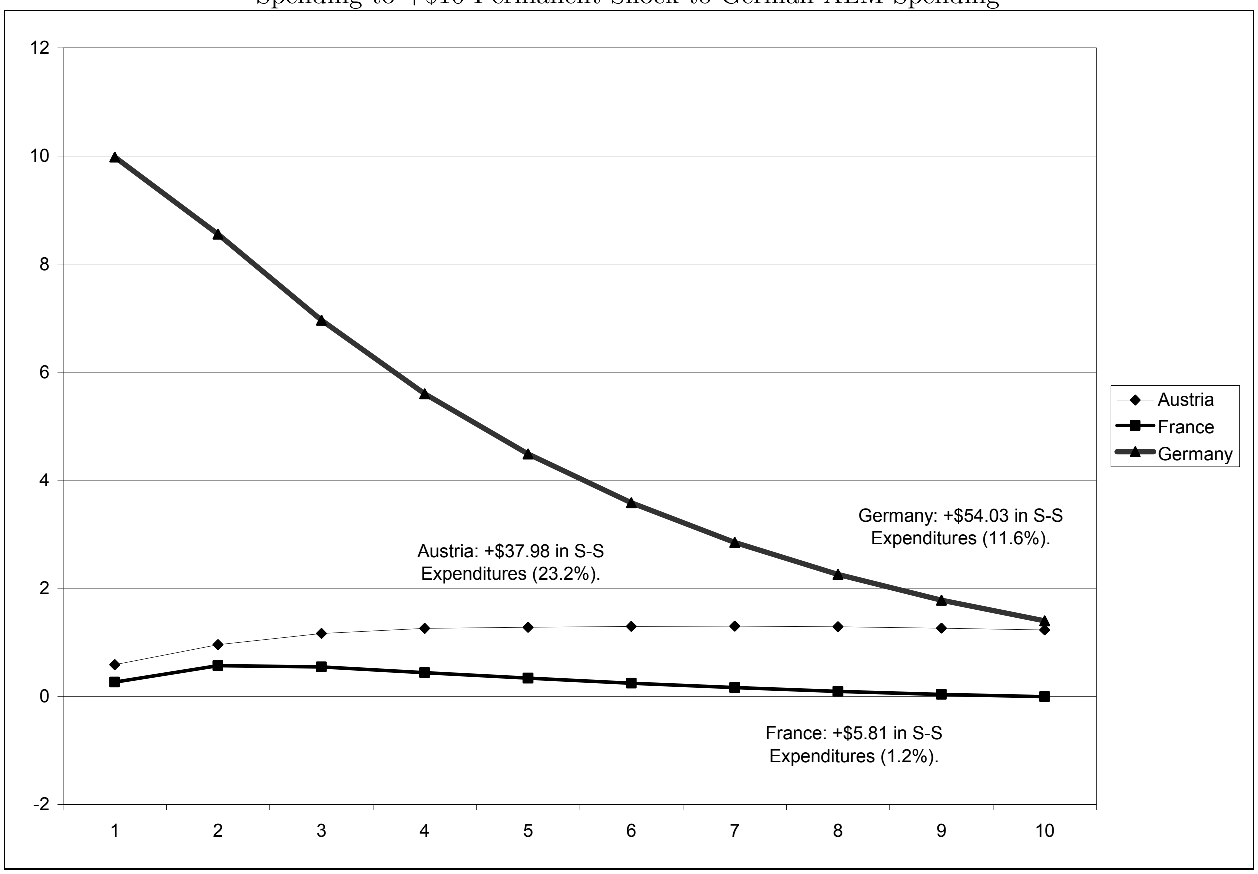
Spatial multipliers, here: $\left[\mathbf{I}_n - \mathbf{W}_t + \gamma\left[abs\left(\mathbf{\Pi}\left[\mathbf{y}_{t-1} \otimes \mathbf{I}_N\right]\right)\right]\right]^{-1}$, capture the feedback from, say, Belgium onto France and other countries, and back from France and those others onto Belgium, and so on. The immediate time-t effect on the vector of policies in the 21 countries, $\mathbf{y}_t$, from a given set of time t-1 policies, $\mathbf{y}_{t\text{-}1}$, including the spatial feedback, can be calculated with this spatial multiplier for any desired counterfactual shocks to the rest of the right-hand side of (37). To find the long-run, steady-state, equilibrium (cumulative) level of $\mathbf{y}$, we must solve (37) recursively. With exogenously time-varying $\mathbf{W}$, like our trade weights, $\mathbf{W}_3$, we need to specify values or sequence of values of $\mathbf{W}$ that we will assume to maintain in the long-run or obtain over the period in question. With endogenously time-varying $\mathbf{W}$, like our policy-distance matrix $\mathbf{W}_4$, the system is much more complex. (For instance, stationarity must not only obtain initially but also hold continually as dynamics unfold.) To get variance estimates, we could use the delta method again, or simulate them by a parametric bootstrap as described above. Given the considerable nonlinearity of (37), the simulated standard-error estimates may have better properties.

Table 5: Effects of a Common Counterfactual ($1) Shock to ALM Spending

	Pre-Shock Steady-State ALM-Spending ($)	*Post-Shock Steady-State ALM-Spending ($)*	*Difference in Steady-State ALM-Spending ($)*
Australia	528.08	525.95	-2.13
Austria	150.52	150.29	-0.23
Belgium	450.67	446.25	-4.42
Canada	156.81	173.08	16.27
Denmark	545.71	544.77	-0.94
Finland	104.48	113.76	9.28
France	478.34	474.95	-3.39
Germany	466.87	465.8	-1.07
Greece	70.11	79.13	9.02
Ireland	85.04	92.6	7.56
Italy	109.87	115.55	5.68
Japan	472.63	476.32	3.69
Netherlands	456.02	452.09	-3.93
New Zealand	120.55	132.16	11.61
Norway	525.74	523.6	-2.14
Portugal	432.16	416.63	-15.53
Spain	100.25	110.89	10.64
Sweden	577.26	575.66	-1.6
Switzerland	332.48	357.31	24.83
United Kingdom	118.28	128.2	9.92
United States	204.54	211.22	6.68

Table 5 illustrates the calculation of estimated coevolutionary spatiotemporal responses to hypothetical shocks. In this example, we start with the 2001 values for all the exogenous variables (the last year all countries have data) and, using the parameter estimates from model 3 of Table 4, determine the steady-state levels of ALM expenditures by recursive calculation of (35). Then, with the system at this steady-state, we shock each country's ALM spending by \$1 (i.e., a \$1 shock to $\boldsymbol{\delta}$) and calculate the new steady-state that emerges from there by the same recursive calculations. The results in Table 5 show the estimated system reasonably stable and the effects sizable. The pattern suggests convergence, with previously low (high) spenders spending more (less) in the new steady state.

<u>Figure 2</u>: Coevolutionary Spatiotemporally Dynamic Response of German, Austrian, and French ALM-Spending to +\$10 Permanent Shock to German ALM-Spending



We can also use (35) to plot estimated coevolutionary spatiotemporal responses to shocks. Using the 2001 values of the exogenous variables, and starting from the steady states that would emerge from those values and the parameter estimates, Figure 2 plots, as an example, the 10-period responses in German, Austrian, and French ALM spending to a \$10 permanent positive shock to Germany's ALM spending ($\boldsymbol{\delta}$). Note the differences in how

France and Austria respond to the German shock, Austria converging toward Germany's permanently higher ALM-spending and France returning to its *status quo ante*. These reflect the latent structure of interdependence between these three countries, as seen in the estimated network arranged in Table 6 and depicted in Figure 3 next.

Using (32), we can also show the estimated weights matrix, i.e., estimated pattern of network interdependencies in ALM policy among these countries, as in Table 6 for the estimated network using the 1981 covariate values.[44]

Table 6: Estimated ALM-Policy Interdependencies, i.e., the Net ALM-Policy Network, in 1981

	AUS	**CAN**	**FIN**	**FRA**	**NTH**	**NWZ**	**ESP**	**SWE**	**GBR**	**USA**
AUS	0.000	0.022	-0.371***	0.160**	-0.182*	-0.905***	-0.025	-0.481**	0.179	1.123**
	(0.000)	(0.095)	(0.140)	(0.075)	(0.096)	(0.326)	(0.023)	(0.195)	(0.200)	(0.566)
CAN	-0.177**	0.000	-0.351***	-0.164**	-0.129**	-0.097***	-0.120***	-0.553***	0.020	1.092
	(0.073)	(0.000)	(0.126)	(0.075)	(0.055)	(0.036)	(0.046)	(0.200)	(0.064)	(0.787)
FIN	-0.283**	-0.159**	0.000	-0.072	0.080	-0.140***	-0.227**	-0.246	0.493	0.075
	(0.114)	(0.075)	(0.000)	(0.165)	(0.102)	(0.051)	(0.101)	(0.317)	(0.308)	(0.180)
FRA	0.047**	-0.071	-0.347**	0.000	-0.033	-0.207***	-0.275*	-0.418**	-0.345	0.565*
	(0.022)	(0.062)	(0.142)	(0.000)	(0.250)	(0.076)	(0.145)	(0.204)	(0.222)	(0.306)
NTH	-0.293**	-0.082	-0.134*	0.051	0.000	-0.039**	-0.175*	-0.238	-0.531*	0.357
	(0.114)	(0.056)	(0.070)	(0.313)	(0.000)	(0.016)	(0.099)	(0.149)	(0.290)	(0.280)
NWZ	-0.463	0.001	-0.244***	-0.220*	-0.005	0.000	-0.214***	-0.439***	0.575**	0.529
	(0.328)	(0.062)	(0.089)	(0.116)	(0.028)	(0.000)	(0.082)	(0.163)	(0.272)	(0.343)
ESP	-0.038	-0.055	-0.381***	-0.220	-0.053	-0.190***	0.000	-0.502**	0.240	0.720**
	(0.025)	(0.047)	(0.146)	(0.308)	(0.124)	(0.069)	(0.000)	(0.214)	(0.215)	(0.355)
SWE	-0.243**	-0.148*	-0.638**	0.033	0.113	-0.169***	-0.190*	0.000	0.560	0.203
	(0.106)	(0.083)	(0.299)	(0.188)	(0.141)	(0.062)	(0.101)	(0.000)	(0.354)	(0.231)
GBR	-0.215*	0.057	-0.175*	-0.619***	-0.405**	0.034*	-0.124	-0.270	0.000	0.633
	(0.114)	(0.080)	(0.104)	(0.218)	(0.186)	(0.021)	(0.094)	(0.198)	(0.000)	(0.387)
USA	0.008	0.284	-0.382***	0.108	-0.069	-0.146**	0.040	-0.54**0	0.217	0.000
	(0.070)	(0.455)	(0.143)	(0.111)	(0.106)	(0.062)	(0.042)	(0.214)	(0.190)	(0.000)

Note: Dependent variable: Total ALM Spending; Model: 3. Actual weights multiplied by 10 (and standard errors adjusted accordingly) to improve table formatting.

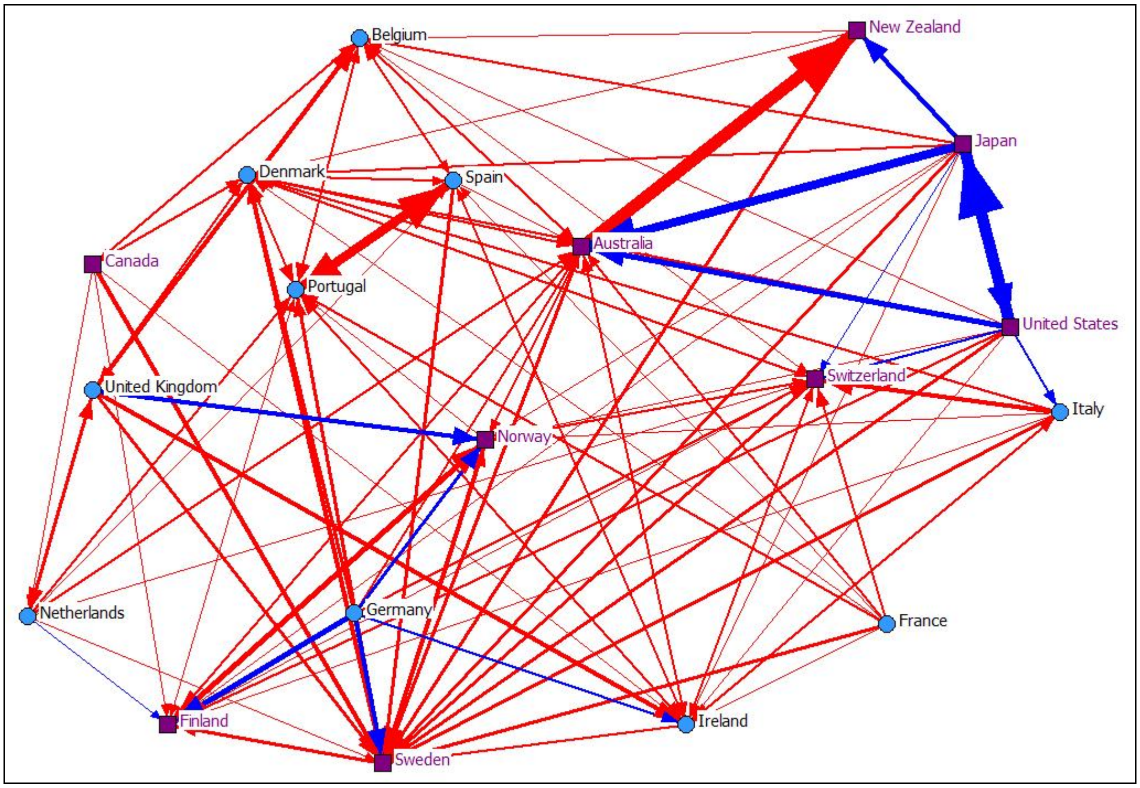
We can also illustrate our estimated patterns of interdependence, i.e., the ALM-policy network, using graphical techniques familiar to network analysts. Figure 4 graphs the estimated patterns and strengths of interdependence in 1991 thus.[45] EU member-countries are circles; other countries are squares. Red arrows represent negative interdependence, or implicit (net) network ties. Blue arrows indicate positive implicit (net) network-ties. Arrow

[44] Tables for the 1991 and 2001 values are too large to show effectively but are available on request.

[45] The web appendix shows the analogous graphs using the 1981 and 2001 covariate values.

thickness and arrow-head sizes gauge the estimated strength of the relationship. We would plot nodes with all estimated dependencies insignificant, say at .10, or of negligible strength, say less than .01, as singletons, but there are no such in this example.

<u>Figure 4</u>: The Estimated Network of ALM-policy Interdependence, 1991



VI. Conclusion and Discussion

In Franzese & Hays (2006c), we estimated single-lag STAR models of ALM policy using binary contiguity (borders) weights matrices and a sample of European countries over the period 1987-1998. Our estimated coefficients on the spatial lags in those regressions were negative and statistically significant, and we argued that these results suggested appreciable ALM-policy free-riding in the EU. The results here, using an m-STAR model to consider multiple possible patterns and pathways of ALM-policy interdependence among the developed democracies more broadly, are strongly consistent with the conclusion that free-riding dynamics dominate among EU members and that these dynamics emerge specifically in great extent due to cross-border spillovers as we had suggested. We also find now some evidence of positive dependence deriving from trade-related competition, supporting globalization-induced competitive-races (not necessarily to bottom) arguments, and that policymakers follow most closely foreign policymakers from similar countries, at least where similarity is gauged by the magnitude of the policy in question. Methodologically, we have offered a simple way to model and estimate networks/interdependence-patterns simultaneously with estimation of the effect of those networks/interdependencies on units' actions. Within this framework, we have suggested and started on the more ambitious agenda of endogenizing those two components of the coevolution of unit behavior/actions and networks/interdependence-patterns.

References

Allers, M., Elhorst, J.P. 2005. "Tax Mimicking and Yardstick Competition among Local Governments in the Netherlands," unpublished manuscript, University of Groningen.

Anselin, L. 2006. "Spatial Econometrics." In T.C. Mills and K. Patterson, eds., *Palgrave Handbook of Econometrics: Volume 1, Econometric Theory*. Basingstoke: Palgrave Macmillan, pp. 901-941.

Anselin, L. 2002. "Under the hood. Issues in the specification and interpretation of spatial regression models." *Agricultural Economics* 27(3):247-67.

Anselin, L. 2003. "Spatial Externalities, Spatial Multipliers and Spatial Econometrics," *International Regional Science Review* 26(2):153-66.

Beck, N., Gleditsch, K.S., Beardsley, K. 2006. "Space is more than geography: Using spatial econometrics in the study of political economy." *International Studies Quarterly* 50:27-44.

Beck, N., Katz, J. 1995. "What To Do (and Not to Do) with Time-Series-Cross-Section Data in Comparative Politics," *American Political Science Review* 89(3):634-47.

Beck, N., Katz, J. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models," *Political Analysis* 6:1-36.

Brueckner, J.K. 2003. "Strategic Interaction Among Governments: An Overview of Empirical Studies," *International Regional Science Review* 26(2):175-188.

Brueckner, J.K., Saavedra, L.A. 2001. "Do Local Governments Engage in Strategic Property-Tax Competition?" *National Tax Journal* 54(2):203-229.

Cameron, D. 1978. "The Expansion of the Public Economy: A Comparative Analysis," *American Political Science Review* 72(4):1243-61.

Case, A.C., Rosen, H.S., Hines, J.R. 1993. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States," *Journal of Public Economics* 52(3):285-307.

Elhorst, J.P. 2001. "Dynamic Models in Space and Time." *Geographical Analysis* 33:119–140.

Elhorst, J.P. 2003. "Specification and Estimation of Spatial Panel Data Models." *International Regional Science Review* 26:244-68.

Elhorst, J.P. 2005. "Unconditional Maximum Likelihood Estimation of Linear and Log-Linear Dynamic Models for Spatial Panels." *Geographical Analysis* 37:85-106.

Florax, R.J.G.M., Folmer, H., Rey, S.J. 2003. "Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology." *Regional Science and Urban Economics* 33(5):557-79.

Florax, R.J.G.M., Folmer, H., Rey, S.J. 2006. "A Comment on Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology: A Reply," *Regional Science and Urban Economics* 36(2):300-8.

Franzese, R., Hays, J. 2003. "Modeling Spatial Relationships in International and Comparative Political Economy: An Application to Globalization and Capital Taxation in Developed Democracies," presented at the Annual Meetings of the Midwest Political Science Association.

Franzese, R., Hays, J. 2004a. "Empirical Modeling Strategies for Spatial Interdependence: Omitted-Variable vs. Simultaneity Biases," presented at the 21st Summer Meeting of the Society for Political Methodology.

Franzese, R., Hays, J. 2004b. "Modeling International Diffusion: Inferential Benefits and Methodological Challenges, with an Application to International Tax Competition," *Wissenschaftszentrum-Berlin SP II 2004 – 12*.

Franzese, R., Hays, J. 2005a. "Modeling Spatial Interdependence in Comparative and International Political Economy with an Application to Capital Taxation," presented at the annual Meeting of the Midwest Political Science Association.

Franzese, R., Hays, J. 2005b. "Spatial Econometric Modeling, with Application to Employment Spillovers and Active-Labor-Market Policies in the European Union," presented at Groningen University, workshop on 'Partisan Politics, Political Autonomy, and Policy Harmonization across Europe'.

Franzese, R., Hays, J. 2006a. “Calculating and Presenting Spatio-Temporal Patterns of Diffusion and Interdependence,” presented at the General Conference of the European Consortium of Political Research, Pisa.

Franzese, R., Hays, J. 2006b. “Spatiotemporal Models for Political-Science Panel and Time-Series-Cross-Section Data,” presented at the 23rd Annual Summer Meetings of the Society for Political Methodology.

Franzese, R., Hays, J. 2006c. “Strategic Interaction among EU Governments in Active-Labor-Market Policymaking: Subsidiarity and Policy Coordination under the European Employment Strategy,” *European Union Politics* 7(2):167-89.

Franzese, R., Hays, J. 2007a. “Empirical Models of International Capital-Tax Competition,” in *International Taxation Handbook*, G. Gregoriou, C. Read, eds., Elsevier Press, pp. 43-72.

Franzese, R., Hays, J. 2007b. “Interdependence in Comparative & International Political Economy, with Applications to Economic Integration and Strategic Fiscal-Policy Interdependence,” presented at *Paris 13 (Université Paris), Axe 5: PSE.*

Franzese, R., Hays, J. 2007c. “Spatial-Econometric Models of Cross-Sectional Interdependence in Political-Science Panel and Time-Series-Cross-Section Data,” *Political Analysis* 15(2):140-64.

Franzese, R., Hays, J. 2007d. *Spatial Econometric Models of Interdependence*. Book prospectus.

Franzese, R., Hays, J. 2008a. “Empirical Modeling of Spatial Interdependence in Time-Series Cross-Sections,” in S. Pickel, G. Pickel, H-J. Lauth, D. Jahn, eds., *Neuere Entwicklungen und Anwendungen auf dem Gebiet der Methoden der vergleichenden Politikwissenschaft, Band II. Wiesbaden: Westdeutscher Verlag*, forthcoming.

Franzese, R., Hays, J. 2008b. “Empirical Models of Spatial Interdependence,” in J. Box-Steffensmeier, H. Brady, D. Collier, eds., *Oxford Handbook of Political Methodology*, Oxford UP.

Franzese, R., Hays, J. 2008c. “Interdependence in Comparative Politics: Theoretical- & Empirical-Model Specification, Estimation, Interpretation, and Presentation,” *Comparative Political Studies* 41(4/5):742-80.

Franzese, R., Hays, J. 2008d. “The Spatial Probit Model of Interdependent Binary Outcomes: Estimation, Interpretation, and Presentation,” presented at the 24th Summer Meeting of the Society for Political Methodology (2007); updates presented at the 2008 Midwest Political Science Association and the 2009 Public Choice Society meetings.

Fredriksson, P.G., Millimet, D.L. 2002. “Strategic Interaction and the Determination of Environmental Policy across U.S. States,” *Journal of Urban Economics* 51(1):101-122.

Garrett, G. 1998. *Partisan Politics in the Global Economy.* New York: Cambridge UP.

Hays, J., Ehrlich, S., Peinhardt, C. 2005. “Government Spending and Public Support for Trade in the OECD: An Empirical Test of the Embedded Liberalism Thesis,” *International Organization* 59(2):473-494.

Hays, J. 2009. “Bucking the System: Using Simulation Methods to Estimate and Analyze Systems of Equations with Qualitative and Limited Dependent Variables.” Paper given at the 2nd Annual St. Louis Area Methods Meeting (SLAMM), Washington University in St. Louis (http://slamm.wustl.edu).

Hays, J., Kachi, A. 2008. “Estimating Interdependent Duration Models with an Application to Government Formation and Survival,” Plenary Session Paper given at the Society for Political Methodology’s 25th Annual Summer Conference, University of Michigan, Ann Arbor (http://polmeth.wustl.edu/conferences/methods2008/program.html).

Hendry, D.F. 2006. “A comment on ‘Specification searches in spatial econometrics: The relevance of Hendry’s methodology’,” *Regional Science and Urban Economics* 36(2):309-12.

Iversen, T., Cusack, T.R. 2000. “The Causes of Welfare State Expansion: Deindustrialization or Globalization?” *World Politics* 52(3):313-49.

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

King, G., Honaker, J., Joseph, A., Scheve, K. 2001. “Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation.” *American Political Science Review* 95(1):49-69.

Kitschelt, H. 1994. *The Transformation of European Social Democracy.* Cambridge UP.

Leenders, Roger T.A.J. 1995. *Structure and Influence*. Dissertation. ICS / University of Groningen. ISBN 90-5170-329-5.

———. 1997. "Longitudinal Behavior of Network Structures and Actor Attributes: Modeling Interdependence of Contagion and Selection," in P. Doreian & F. Stokman, eds., *Evolution of Social Networks*, pp. 165-84. Amsterdam: Routledge.

LeSage, J.P. 1999. *Spatial Econometrics*. rri.wvu.edu/WebBook/LeSage/spatial/spatial.html.

Redoano, M. 2003. "Fiscal Interactions among European Countries," *Warwick Economic Research Papers* No. 680.

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

Ruggie, J.G. 1982. "International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order," *International Organization* 36(2):195-231.

Simmons, B., Elkins, Z. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy," *American Political Science Review* 98(1): 171-190.

Snijders, T.A.B. 1997. "Stochastic Actor-Oriented Models for Network Change," in P. Doreian & F. Stokman, eds., *Evolution of Social Networks*, pp. 185-208. Amsterdam: Routledge.

———. 2001. The Statistical Evaluation of Social Network Dynamics. In M.E. Sobel and M.P. Becker, eds., *Sociological Methodology* 31:361-95.

———. 2005. Models for Longitudinal Network Data. Chapter 11 in P. Carrington, J. Scott and S. Wasserman, eds., *Models and methods in social network analysis*. New York, Cambridge UP.

Snijders, T.A.B., Borgatti, S.P. 1999. "Non-Parametric Standard Errors and Tests for Network Statistics," *Connections* 22(2):161-70.

Snijders, T.A.B., Steglich, C., Schweinberger, M. 2007. Modeling the co-evolution of networks and behavior. Chapter 3 in K. van Montfort, H. Oud and A. Satorra, eds., *Longitudinal models in the behavioral and related sciences*. Mahwah NJ, Lawrence Erlbaum. pp.41-71.

Snijders, T.A.B., Steglich, C., West, P. 2006. "Applying SIENA: An illustrative analysis of the co-evolution of adolescents' friendship networks, taste in music, and alcohol consumption."

Swank, D. 2002. *Global Capital, Political Institutions, and Policy Change in Developed Welfare States* Cambridge: Cambridge UP.