ABSTRACT: Spatial interdependence, the interdependence of outcomes across units, is theoretically and substantively ubiquitous and central across the social sciences. The empirical association, correlation, or clustering 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 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 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). We again 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, 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, EU co-membership, and geographic proximity 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 Darnton, Jakob de Haan, Scott De Marchi, John Dinardo, Zach Elkins, James Fowler, John Freeman, Fabrizio Gilardi, Kristian Gleditsch, Mark Hallberg, John Jackson, Jonathan Katz, Mark Kayser, Achim Kenmerling, Gary King, Greg Kojer; Hasan Kirmanoglu, Herbert Kitschelt, Jim Kuklinski, Don Lacombe, Chris Lawrence, David Lazer, Tse-Min Lin, Xiaoobo Lu, Scott McClurg, Walter Mebane, Covadonga Meseguer, Mark Mizruchi, Michael Peress, Thomas Pfemper, David Prosperi, Dennis Quian, Megan Reif, Frances Rosenbluth, Ken Scheve, John Scholz, Phil Schrodt, Ken Schultz, Jasjeeb Sekhon, Chuck Shipan, Beth Simmons, Betsy Sinclair, David Siroky, John D. Stephens, Duane Swank, Wendy Tam-Cho, Vera Troeger, Craig Volden, Michael Ward, Greg Wa waro, 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.

** Work addressing fully simultaneous endogeneity of selection and comparison of network-analytic and spatial-econometric approaches still in progress.

KEYWORDS: Interdependence, Spatial Lag, Spatial Econometrics, m-STAR: Multiparametric Spatiotemporal Autoregressive Model, Network Dependence, Coevolution.
I. INTRODUCTION: 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-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

1 The ensuing list of topics and disciplines corresponds to literature searches for applied work under contagion, spatial interdependence, or network dependence. An 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.
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 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).

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.
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 \((p_i, p_j)\):

\[
V^i \equiv V^i(p_i, p_j) ; \quad V^j \equiv V^j(p_j, p_i)
\]  

(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:

\[
p^*_i = \text{Argmax}_{p_i} V^i(p_i, p_j) = R(p_j) ; \quad p^*_j = \text{Argmax}_{p_j} V^j(p_i, p_j) = R(p_i)
\]  

(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-partial:

\[
\frac{\partial p^*_i}{\partial p_j} = -\frac{V^i_{p_i p_j}}{V^i_{p_j p_j}} ; \quad \frac{\partial p^*_j}{\partial p_i} = -\frac{V^j_{p_j p_i}}{V^j_{p_j p_j}}
\]  

(3).

If governments are maximizing, the denominators are negative, so, if \(V^{ij}_{pp} > 0\), policies are strategic complements: reaction-functions slope upward. If \(V^{ij}_{pp} < 0\), reaction functions slope downward: policies are strategic substitutes. If \(V^{ij}_{pp} = 0\), 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.\(^4\) 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

\(^4\) In Europe at least, these apparently strong labor-market externalities along borders seem empirically to owe much more to firm-location decisions than to any actual labor mobility (see, e.g., Overman & Puga 2002).
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).\footnote{To give another example (from Masket 2008): representatives who sit together may vote together because they sit by party and so have similar constituencies, or because they talk and influence each other, or they may sit together because they know and like each other which may be for the same reasons they vote similarly.}

We discussed elsewhere (Franzese & Hays 2003, 2004ab, 2006abc, 2007abcd, 2008abc) the severe empirical-methodological challenges in estimating interdependence and contagion distinctly and well (a.k.a, Galton’s Problem\footnote{The web appendix contains, inter alia, brief intellectual-historical background to the label.}). 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 are therefore 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 extension offered (along with an introduction and discussion of network-analytic models of coevolution) 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 analytic results comparing the asymptotic properties, and simulation evidence comparing the small-sample performance, of full-information spatial maximum-likelihood (S-ML) estimates of mSTAR models in these contexts to those of a naïve estimator applying least-squares to a linear regression including multiple spatial-lags and to those of a blind estimator that applies least-squares omitting spatial lags. Section V illustrates this approach to dynamic, endogenous interdependence with an empirical application attempting to disentangle the roles of economic interdependence, correlated external and internal stimuli, EU co-membership, and geographic proximity 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, 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:

\[ y_{it} = \rho \sum_{j \neq i} w_{ij} y_{jt} + \phi y_{i,t-1} + \beta_d' d_{it} + \beta_s' s_{it} + \beta_{sd}' \left( d_{it} \odot s_{it} \right) + \varepsilon_{it} \]  

(4.7)

\( y_{jt} \) is the outcome in another \((j \neq i)\) unit, which in some manner (given by \( \rho w_{ij} \)) directly affects the outcome in unit \( i \). The \( w_{ij} \) reflect the relative connectivity from \( j \) to \( i \), and \( \rho \) 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 co-membership of \( i \)'s and \( j \)'s political economies. The other right-hand-side factors reflect the non-interdependence components: unit-level/domestic factors, \( d_{it} \) (e.g., election-year indicators, government partisanship), exogenous-external or -contextual factors, \( s_{it} \) (e.g., technology, oil prices; merely for contrast, assume these common

\footnote{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.}
across units: $s_j$), and context-conditional factors, $d_{it} \odot s_i$ (i.e., the interactions of the former with the latter). The $\varepsilon_{it}$ are i.i.d. stochastic terms.8

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.9 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, $w_{ij}$, 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, intuitively, inadequacies or omissions in specifying the one tend to induce underestimates of that one’s role and overestimates of the other’s role.

Secondarily,10 even with the interdependence and the alternative common-shock mechanisms modeled perfectly, the spatial-lag regressor(s) will be endogenous (i.e., covary with $\varepsilon$), so estimates of $\rho$ 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 (typically overestimation) generally induce biases in the opposite direction (typically 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

---

8 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.

9 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).

10 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.
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:11

\[
\mathbf{y} = \rho \mathbf{Wy} + \phi \mathbf{My} + \mathbf{X}\beta + \mathbf{\varepsilon} \tag{5}
\]

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 \( T \)).12 \( \rho \) is the previously described spatial-autoregressive coefficient, and \( \mathbf{W}_{NT} \) is an \( NT \times NT \) block-diagonal spatial-weighting matrix.13 Each of the \( T \) \( N \times N \) weights matrices, \( \mathbf{W}_t \), on the block-diagonal have elements \( w_{ij}(t) \) reflecting the relative connectivity from unit \( j \) to \( i \) that period.14 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 \((N+1,1),(N+2,2),\ldots, (NT,NT-1)\), and zeros elsewhere. \( \mathbf{My} \) is thus a standard (first-order) temporal-lag;15 \( \phi \) is its coefficient. \( \mathbf{X} \) contains \( NT \) observations on \( k \) independent variables; \( \beta \) is its \( k \times 1 \) vector of coefficients, and \( \mathbf{\varepsilon} \) is an \( NT \times 1 \) vector of i.i.d. stochastic components.16

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), which is 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 S-OLS produce biased and inconsistent estimates, OLS due to the omitted-variable bias and S-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

---

11 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.

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

13 \( \mathbf{W}_{NT} \) 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.

14 If the pattern of connectivity is time-invariant, then \( \mathbf{W}_{NT} \) can be expressed as the Kronecker product of a \( TxT \) identity matrix and the constant \( NxN \) weights-matrix, \( \mathbf{I}_T \otimes \mathbf{W}_N \).

15 Higher-order time-dynamics would simply add further properly configured weights matrices.

16 Again, alternative distributions of \( \mathbf{\varepsilon} \) are possible but add complication without illumination.
interdependence, OLS coefficient-estimates will suffer omitted-variable biases—familiarly: $F \beta$, where $F$ is the matrix of coefficients obtained by regressing the omitted on the included variables and $\beta$ is the vector of (true) coefficients on the omitted variables.\(^{17}\) In this case, the omitted-variable bias (OVB) is:

$$
\text{OVB}[\hat{\beta}_{\text{OLS}}, \hat{\phi}_{\text{OLS}}]' = \rho \times (Q_i Q_i)^{-1} Q_i' W_y, \quad \text{where } Q_i = \begin{bmatrix} X & M_y \end{bmatrix}'
$$

(6).

$\rho_{\text{OLS}} = 0$, 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, (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, $W_y$, covaries with the residual, $\varepsilon$, 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:

$$
\text{SB}[\hat{\rho}, \hat{\phi}, \hat{\beta}] = (Q'Q)^{-1} Q' \varepsilon, \quad \text{where } Q = \begin{bmatrix} W_y & M_y & X \end{bmatrix}'
$$

(7).

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

$$
\text{SB}[\hat{\rho}, \hat{\phi}, \hat{\beta}] = \frac{1}{\Psi} \begin{bmatrix} \text{Cov}(W_y, \varepsilon) \times \text{Var}(M_y) \times \text{Var}(x) \\ -\text{Cov}(W_y, \varepsilon) \times \text{Cov}(W_y, M_y) \times \text{Var}(x) \\ -\text{Cov}(W_y, \varepsilon) \times \text{Cov}(W_y, x) \times \text{Var}(M_y) \end{bmatrix}, \quad \text{where } \Psi = \text{plim} \left( \frac{Q'Q}{n} \right)
$$

(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, \textit{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 $W$, 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, exogenous-external, and context-conditional effects. Again, those factors that correlate most with the interdependence pattern will have the most severe induced deflation biases. These

\(^{17}\) Estimates of limited- or qualitative-dependent-variable models, like logit or probit, which exclude relevant spatial lags will suffer analogous omitted-variable biases.
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 \( W \) 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, not only to scholars working at those levels, but also to those interested in gauging the role of interdependence.

Our simulations (Franzese & Hays 2004a, 2006b, 2007cd) showed the omitted-variable biases of OLS are usually worse and often far, far worse than S-OLS’ simultaneity biases. In fact, S-OLS may perform adequately for mild interdependence strengths \( (\rho \Sigma_w \leq .25) \), although standard-error accuracy can be problematic, and in a manner for which PCSE (Beck & Katz 1995, 1996) will not compensate. However, S-OLS’ simultaneity biases do grow sizable as interdependence strengthens, 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,\(^{18}\) we introduce only it here.\(^{19,20}\)

The conditional likelihood function for the spatiotemporal-lag model,\(^{21}\) 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:

\(^{18}\) 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 MatLab 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 \( x \)'s in others), \( WX \) are ideal instruments by construction. Cross-spatial endogeneity may seem unlikely, until one realizes that vertical ties \( y_i \) to \( y_j \) (interdependence) and horizontal ties from \( y_j \) to \( x_j \) (typical simultaneity) combine to give the offending diagonals from \( y_i \) to \( x_j \). Nonetheless, 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.

\(^{19}\) Initially, we used J.P. LeSage’s MatLab\(^TM\) code to estimate our spatial models, having found early third-party contributed Stata\(^TM\) code for spatial analysis untrustworthy and/or extremely computer-time intensive. We have since written MatLab\(^TM\) code to implement all, and Stata\(^TM\) code to implement many, of our suggestions. For code, plus Excel\(^TM\) spreadsheets useful as templates for interpretation and presentation: https://netfiles.uiuc.edu/jchays/www/page.html and http://www.umich.edu/~franzese/Publications.html.

\(^{20}\) Franzese & Hays (2008d), Hays & Kachi (2008), and Hays (2009) introduce for political science empirical models of spatial interdependence in limited and qualitative dependent-variables.

\(^{21}\) Derivation of likelihoods for the spatiotemporal-lag models is due to Elhorst (2001, 2003, 2005).
\[ \mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \mathbf{\beta} + \mathbf{\varepsilon} \Rightarrow \mathbf{\varepsilon} = (I - \rho \mathbf{W}) \mathbf{y} - \mathbf{X} \mathbf{\beta} \equiv \mathbf{A} \mathbf{y} - \mathbf{X} \mathbf{\beta} \] (9),

where \( \mathbf{X} \) now includes \( \mathbf{M} \mathbf{y} \), the time-lag of \( \mathbf{y} \), as its first column, and \( \mathbf{\beta} \) includes \( \phi \) as its first row.22 Assuming i.i.d. normality, the likelihood function for \( \mathbf{\varepsilon} \) is the typical linear-normal:

\[ L(\mathbf{\varepsilon}) = \left( \frac{1}{\sigma^2 2\pi} \right)^{\frac{NT}{2}} \exp \left\{ -\frac{\mathbf{\varepsilon}' \mathbf{\varepsilon}}{2\sigma^2} \right\} \] (10),

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

\[ L(\mathbf{y}) = |\mathbf{A}| \left( \frac{1}{\sigma^2 2\pi} \right)^{\frac{NT}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{A} \mathbf{y} - \mathbf{X} \mathbf{\beta})'(\mathbf{A} \mathbf{y} - \mathbf{X} \mathbf{\beta}) \right\} \] (11).

This resembles the typical linear-normal likelihood, except the transformation from \( \mathbf{\varepsilon} \) to \( \mathbf{y} \) is not by the usual factor, 1, but by \( |\mathbf{A}| = |I - \rho \mathbf{W}| \).23 Written in \((N \times 1)\) vector notation, the spatiotemporal-model conditional-likelihood is mostly conveniently separable into parts:

\[ \log f_{y_1, y_2, \ldots, y_T} = -\frac{1}{2} N (T - 1) \log(2\pi \sigma^2) + (T - 1) \log |\mathbf{A}| - \frac{1}{2\sigma^2} \sum_{t=2}^{T} \mathbf{\varepsilon}_t^{\prime} \mathbf{\varepsilon}_t \] (12).

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

\[ \log f_{y_1, y_2, \ldots, y_T} = -\frac{1}{2} N (T - 1) \times \log(2\pi \sigma^2) + (T - 1) \log |\mathbf{A}| - \frac{1}{2\sigma^2} \sum_{t=2}^{T} \mathbf{\varepsilon}_t^{\prime} \mathbf{\varepsilon}_t \]

\[ + \frac{1}{2} \sum_{i=1}^{N} \log \left( 1 - \rho_i \right)^2 - \frac{1}{2\sigma^2} \mathbf{\varepsilon}_1^{\prime} (\mathbf{A} - \phi \mathbf{I}_N)^{-1} (\mathbf{A}^{\prime} \mathbf{A}^{-1} (\mathbf{A}^{\prime} \mathbf{A}^{-1})^\prime) (\mathbf{A} - \phi \mathbf{I}_N)^{-1} \mathbf{\varepsilon}_1 \] (13).

With large \( T \), the contribution of the first observation to the total likelihood—namely, the last three terms—is relatively little, so the simpler conditional likelihood can serve adequately.

One easy way to ease or even erase S-OLS’s simultaneity problem is to lag temporally the spatial lag (as, e.g., Beck et al. 2006 suggest). 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:

\[ \mathbf{y}_t = \eta \mathbf{W} \mathbf{y}_{t-1} + \phi_1 \mathbf{y}_{t-1} + \mathbf{X} \mathbf{\beta} + \mathbf{\varepsilon}_t \] (14).

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

\[ \log f_{y_1, y_2, \ldots, y_T} = -\frac{1}{2} NT \times \log(2\pi \sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^{T} \mathbf{\varepsilon}_t^{\prime} \mathbf{\varepsilon}_t \]

\[ + \frac{1}{2} \sum_{i=1}^{N} \log \left( 1 - \rho_i \right)^2 - \frac{1}{2\sigma^2} \mathbf{\varepsilon}_1^{\prime} (I - \mathbf{B})^{-1} (I - \mathbf{BB}^\prime)^{-1} (I - \mathbf{B})^{-1} \mathbf{\varepsilon}_1 \] (15),

where \( \mathbf{\varepsilon}_1 = \mathbf{y}_1 - (\phi + \eta \mathbf{W}_N) \mathbf{y}_{t-1} - \mathbf{X} \mathbf{\beta} \), \( \mathbf{\varepsilon}_t = \mathbf{y}_t - \eta \mathbf{W}_N \mathbf{y}_{t-1} - \phi \mathbf{y}_{t-1} - \mathbf{X} \mathbf{\beta} \), and \( \mathbf{B} = \phi \mathbf{I}_N + \eta \mathbf{W}_N \).

---

22 N.b., although \( \mathbf{W} \mathbf{y} \) complicates the conditional likelihood in terms of \( \mathbf{y} \) (see note 23), \( \mathbf{M} \mathbf{y} \) does not.

23 This difference follows from the change-in-variables theorem of basic statistics. This complicates estimation somewhat in that \( |\mathbf{A}| \) involves \( \rho \), so that \( NT \times NT \) determinant requires recalculation at each iteration of the likelihood-maximization routine.
Note that the last two 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 summary, 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 consistent 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.\(^\text{24}\) E.g., a standard one-directional test against spatial-lag alternative is:

\[
LM_\rho = \hat{\sigma}_\epsilon^2 \left( \hat{\epsilon}'W_y / \hat{\sigma}_\epsilon^2 \right) \left[ G + R\hat{\sigma}_\epsilon^2 \right]^{-1}, \quad \text{where} \\
G = (WX\hat{\rho})(I - X'X)^{-1}X'(WX\hat{\rho}) \quad \text{and} \quad R = \text{tr}(W' + W)W
\]

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

\[
LM^*_\rho = G^{-1}\hat{\sigma}_\epsilon^2 \left( \hat{\epsilon}'W_y / \hat{\sigma}_\epsilon^2 - \hat{\epsilon}'W\hat{\epsilon} / \hat{\sigma}_\epsilon^2 \right)^2
\]

(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 \( \omega \) be an eigenvalue of \( W \); then the spatiotemporal process is covariance stationary if:

\[
\left| \phi \left( I - \rho W \right)^{-1} \right| < 1, \quad \text{or, equivalently, if} \\
\begin{align*}
\phi &< 1 - \rho \omega_{\text{max}}, \quad \text{if} \ \rho \geq 0 \\
\phi &< 1 - \rho \omega_{\text{min}}, \quad \text{if} \ \rho < 0
\end{align*}
\]

(18).

For instance, with positive temporal and spatial dependence and \( W \) row-standardized, the \( \omega_{\text{max}} = 1 \), 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 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

\(^{24}\) Be sure to note the corrections posted here: http://www.umich.edu/~franzese/Publications.html.
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\textsuperscript{25} effect of permanent shocks.\textsuperscript{26} 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 \((N\times1)\) vector form:

\[
y_t = \rho W y_t + \phi y_{t-1} + X_t \beta + \epsilon_t
\]

Set \(y_{t-1}\) equal to \(y_t\) fix exogenous RHS terms, \(X\) and/or \(\epsilon\), to their hypothetical permanent post-shock levels, and solve for the long-run steady-state level of \(y\) (assuming stationarity):\textsuperscript{27}

\[
y_t = \rho W y_t + \phi y_{t-1} + X_t \beta + \epsilon_t = (\rho W + \phi I) y_t + X_t \beta + \epsilon_t = [I_N - \rho W - \phi I_N]^{-1} (X_t \beta + \epsilon_t)
\]

Decomposing \(\epsilon_t = \delta + \upsilon_t\) with \(\delta\) fixed and \(\upsilon_t\) 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,\textsuperscript{28} first denote the \(i\)th column of \(S\) as \(s_i\) and its estimate as \(\hat{s}_i\). The steady-state spatiotemporal equilibrium responses to a one-unit increase in the \(i\)th element of \(\delta\) are then \(s_i\), with asymptotic approximate variance-covariance matrix:

\[
\begin{bmatrix}
  1 - \phi & -\rho w_{12} & \ldots & \ldots & -\rho w_{1N} \\
  \vdots & 1 - \phi & \ddots & \ddots & \vdots \\
  -\rho w_{N,1} & \ldots & \ldots & 1 - \phi & -\rho w_{N,(N-1),N}
\end{bmatrix}^{-1}
\]

Here, \(\hat{\theta} = [\hat{\rho}, \hat{\phi}]\), \(\begin{bmatrix} \frac{\partial \hat{s}_i}{\partial \hat{\rho}} & \frac{\partial \hat{s}_i}{\partial \hat{\phi}} \end{bmatrix}\), and the vectors \(\frac{\partial \hat{s}_i}{\partial \hat{\rho}}\) and \(\frac{\partial \hat{s}_i}{\partial \hat{\phi}}\) are the \(i\)th columns of \(\hat{S}\) and of \(\hat{S}\). Similarly, the steady-state spatiotemporal responses to a one-unit increase in explanatory variable \(k\) in unit \(i\) are \(s_i \hat{\beta}_k\), with delta-method standard-errors of

\[
\begin{bmatrix}
  \frac{\partial \hat{s}_i \hat{\beta}_k}{\partial \hat{\theta}} \\
  \frac{\partial \hat{s}_i \hat{\beta}_k}{\partial \hat{\theta}}
\end{bmatrix}
\]

\textsuperscript{25} We use the terms long-run, steady-state, and equilibrium effects loosely, and interchangeably, to refer to the estimated asymptotic level of outcomes \(y\) following a hypothetical permanent shock.

\textsuperscript{26} Anselin (2003) and Franzese & Hays (2006b,2007cd, 2008bcd) discuss these multipliers more fully.

\textsuperscript{27} In the case of time-variant \(W_N\) one must also fix the \(\{w_{ij}(t)\}\) to some desired set or sequence of sets of values.

\textsuperscript{28} We have used only first-order approximations. Higher orders would presumably yield greater accuracy; simulation (e.g., parametric bootstrapping) may also be advantageous.
with $\hat{\beta} = [\hat{\rho} \; \hat{\phi} \; \beta_i]'$ and $
abla_\theta \hat{\beta} = [\nabla_\rho \hat{\beta} \; \nabla_\phi \hat{\beta} \; \nabla_{\theta'} \hat{\beta}_i]$, and with $\nabla_\rho \hat{\beta}$ and $\nabla_\phi \hat{\beta}$ the $i^{th}$ columns of $\hat{\beta}_i S_{WS}$ and $\hat{\beta}_i S_{SS}$, respectively.

One can find the spatiotemporal response path of the $N \times 1$ vector of unit outcomes, $y_t$, to exogenous right-hand-side terms, $X$ and $\varepsilon$, by rearranging (19) to isolate $y_t$ on the left:

$$y_t = [I_N - \rho W]^{-1} \{\phi y_{t-1} + X_\beta + \varepsilon_t\} = S \{\phi y_{t-1} + X_\beta + \varepsilon_t\}$$

This formula gives response-paths of all units to hypothetical shocks to $X$ or $\varepsilon$ in any unit(s) $\{j\}$, including shocks in $\{i\}$ itself/themselves, by setting $(X_\beta + \varepsilon)$ 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 $x$ (cumulative), and their standard errors, working with the $NT \times NT$ matrix is easier. Redefine $S$ in (20) as $S = [I_N - \rho W - \phi M]^{-1}$ 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 of 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 exogenous domestic or unit-level, external, and context-conditional factors. The spatial-ML estimation 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.\(^{29}\) In the network-analysis tradition, Leenders (1995, 1997) and Snijders and colleagues (1997, 2001, 2005, 2007ab) have advanced furthest on this crucial 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’ model (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 the reverse process, in which actors’ attributes shape the network, selection. 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 or edge (i.e., a binary tie) from \( j \) to \( i \) forms, \( w_{ij}=1 \), or dissolves, \( w_{ij}=0 \), at rates, \( \lambda_{ij} \) and \( \lambda_{ji} \), given by some observable attribute(s) of \( i \) and/or \( j \):

\[
\lambda_{ij} = \lambda_{i} + \nu_{i}d_{ij} \quad ; \quad \lambda_{ij} = \lambda_{j} + \nu_{j}d_{ij}
\]

(24),

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

\[
y = \rho W y + X\beta + \epsilon
\]

(25),

which could extend to the standard spatiotemporal model, (5), straightforwardly. Leenders

\(^{29}\) 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.
(1997) integrates these contagion and selection models thus. First, let \( A_t \) be the \( N \times N \) matrix of current realizations of (24), \( y_t \) be the \( N \times 1 \) vector of attributes for the actors, and \( X_t \) the \( N \times k \) matrix of exogenous explanators thereof. Leenders (1997:172) expresses \( W_t = W(A_t) \), 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 as implemented the function is just the identity: \( W_t = A_t \). The model is then identified for estimation of \( \lambda, \nu, \mu, \) and \( \beta \) from \( W_t \) and \( y_t \) observed at discrete intervals \( t=\{1...T\} \) by time lags and the assumptions 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:

\[
W_t = f(W_{t-1}, y_t) \quad w_{ij,t} = d_{ij,t} = |y_{i,t} - y_{j,t}| \quad y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + X_t \beta + \epsilon_t \quad (26).
\]

He then generates \( W_0, y_0, \{\epsilon_t\}, \) and \( \{X_t\} \) randomly, and assesses by simulation the biases entailed in estimating (26) from data collected at intervals of increasing length (measured in 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’\(^{31}\) 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.

\(^{30}\) 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, \( \mu \), interpreted as a societal norm for \( y \). We have simplified to a first-order time-lag to enhance comparability in exposition.

\(^{31}\) We follow Snijders et al. (1997) and Steglich et al. (2006, 2007) most specifically.
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 \( x_{ij}(t) \)—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 \( z_i(t) \)—in our notation, \( \mathbf{y}(t) \), with elements \( y_i(t) \). Further exogenous explanators may exist at unit or dyadic level, \( v_i(t) \) or \( w_{ij}(t) \)—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, \( \lambda_{net} \), according to an exponential hazard-rate model. The model may further parameterize \( \lambda \), but the illustrative 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 \( \lambda_{beh} \).³²

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:

\[
U_{i}^{net} = f_{i}^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z}) + \varepsilon_{i}^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z}) = \sum_{h} \left\{ \beta_{h}^{net} \times s_{h}(i, \mathbf{x}, \mathbf{x}', \mathbf{z}) \right\} + \varepsilon_{i}^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z}) \quad (27),
\]

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 \( i \).³³ \( s_{h}^{net}(\cdot) \) 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 \( \beta_{h}^{net} \) to be estimated are the relative weights of these objectives. Assuming the \( \varepsilon_{i}^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z}) \) extreme-value distributed, independently across options and actors (see note 33) 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.³⁴

As in Leenders’ approach, identification derives from debarring any literal simultaneity in outcomes or networks and assuming that temporal precedence implies causal precedence, and

³² Since observation occurs at discrete intervals, the degrees of freedom to vary these continuous-time rates 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.

³³ 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 35) regarding which if any \( x_{i} \) to switch on or off.

³⁴ 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.
in particular conditioning on the first observation.\textsuperscript{35} Given all this, estimation occurs by simulating the sequences of policies $z$ and of networks $x$ and searching over possible values of the model parameters, $\lambda$ and $\beta$, to minimize some distance function from the observed sequences of $x$ and $z$ to the simulated sequences.\textsuperscript{36} 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.\textsuperscript{37}

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.\textsuperscript{38} 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

\textsuperscript{35} 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. 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 generally include the actual outcomes on the right-hand side, however lagged or transformed by some network-structure measure-function. Only the latent variable or the estimated probabilities can enter those functions. (The problem is that the actual outcomes indirectly enter the right-hand side to predict their own probabilities: see Heckman 1978.) Inclusion of network summary statistics on the right-hand side of these models rests on the $p^*$ and, more generally, ERGM approach to network models. These approaches exploit a property of the exponential-family distributions, which include logit and the exponential, that controlling sufficient statistics for the outcomes yields unbiased estimates of the remaining partial coefficients. Whether the array of possible arguments to the multinomial logits and exponentials in the Snijders et al. model are valid sufficient statistics in this sense seems not assured to us. The presence of a stochastic component exhibiting dependence across units, moreover and most importantly, 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 yet, 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.)

\textsuperscript{36} How exactly the observed frequency is mapped to the simulation “microsteps” is unclear from the exposition.

\textsuperscript{37} The earlier of the works reviewed here indicated that 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.

\textsuperscript{38} This limitation seems not inherent. First, Snijders et al.’s SIENA seems to require only discrete, not necessarily ordinal, behaviors. The limit seems the sensibility of conceiving an option to increment, leave unchanged, or decrement behavior by one. If so, rounding or rescaling continuous behaviors to discrete should suffice, and unbounded behaviors would actually simplify by removing need to alter actors’ choice problem at the bounds.
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 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.\footnote{These obviously are general concerns, not necessarily restricted to C&IPE.}

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,\footnote{The distinction may be seen as implying regressors of the forms $Wy$ in the former case and $f(i,W)$ in the latter.} 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. Starting with the spatio-temporal-lag model, (5), we expand its specification to allow estimation of $W$, 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. For example, the sociologists’ homophily (like seeks or mimics like), if it stems from fixed or exogenous characteristics of ego and alter, parallels from the spatial-econometric perspective a model of $w_{ij}$ as a function of $x_i$ and $x_j$. If we instead consider some function of the vector of behaviors of interest, $y$, among these explanators of $W$, 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 $m$-STAR, a spatiotemporal-lag model with multiple spatial-weights matrices:

$$y = \rho_1 W_1 y + \rho_2 W_2 y + \ldots + \rho_R W_R y + \phi M y + X \beta + \epsilon$$

$$= W y + \phi M y + X \beta + \epsilon,$$

where $W \equiv \sum_{r=1}^{R} \rho_r W_r$ (28).

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

$$y_i = \rho_1 \sum_j w_{ij} y_j + \rho_2 \sum_j w_{ij}^2 y_j + \ldots + \rho_R \sum_j w_{ij}^R y_j + \phi y_{i,t-1} + \sum_k x_k \beta_k + \epsilon_i$$

$$= \sum_j \left( \rho_1 w_{ij}^1 + \rho_2 w_{ij}^2 + \ldots + \rho_R w_{ij}^R \right) y_j + \phi y_{i,t-1} + \sum_k x_k \beta_k + \epsilon_i$$

$$= \sum_j \left( \sum_r \rho_r w_{ij}^r \right) y_j + \phi y_{i,t-1} + \sum_k x_k \beta_k + \epsilon_i$$

(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 $\rho_r$ their coefficients to estimate. Thus, we can conceive the $m$-STAR model as a STAR model with the estimated $W$, i.e., the estimated network, being a weighted sum of observed explanators of connectivity, $\hat{W} = \sum_r \hat{\rho}_r W_r$. If, furthermore, any $W$, has functions of $y$ as elements, then $W$ and $y$ are jointly endogenous, and (29) is a network-behavior coevolution model.

Without considerable further complication, the sorts of models of $W$, 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 a log-odds transformation and a decision rule to convert those log-odds to (1,0). (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. Also, the costs in estimation complexity of enriching the model of connectivity by adding covariates is high, at least compared to adding unit, dyad, or exogenous-external covariate $x$ in $X \beta$ (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-W, case, both for the simpler likelihood conditional on the first observation and for 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 converse side, we can interpret and present the estimated network, \( \hat{W} \), with all the standard network-analytic tools and representations. The next section illustrates both.

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

\[
\ln L(p, \beta, \phi, \sigma; y, X) = \ln(2\pi \sigma^2)^{-NT/2} + \ln|\mathbf{A}| - \frac{1}{2\sigma^2} \mathbf{e}'\mathbf{e},
\]

where \( \mathbf{A} = \mathbf{I}_{NT} - \mathbf{W} \) and \( \mathbf{e} = \mathbf{A}y - \phi \mathbf{My} - \mathbf{X}\beta \) (30).

Written for \((N\times1)\) vectors \( y \), the likelihood is conveniently separable as follows, highlighting the conditionality on the first observation (which is not apparent in (30)):

\[
\log f_{y_1, y_2, \ldots, y_T} = \frac{-1}{2}N(T-1)\log(2\pi\sigma^2) + (T-1)\sum_{i=1}^{T} \log|\mathbf{I} - \rho_i \mathbf{W}_i| - \frac{1}{2\sigma^2}\sum_{i=1}^{T} \mathbf{e}_i'\mathbf{e}_i
\]

where \( \mathbf{e}_i = y_i - \sum_{i=1}^{T} \rho_i W_i y_i - \phi I_y y_{i-1} - \mathbf{X}\beta \) (31).

The unconditional (exact) likelihood extends the more complex (13) analogously. Luckily, \( T \) is large enough in our application, and in most C&IPE and many other contexts, that the more compact conditional likelihood is adequate. In either case, the estimated variance of \( \hat{W} \) is:

\[
\hat{\mathbf{W}} = \sum_{i} \hat{\rho}_i \mathbf{W}_i \Rightarrow \text{Var}(\mathbf{W}^{(ij)}) = \left[ \mathbf{W}^{(i,j)}_1 \mathbf{W}^{(i,j)}_2 \ldots \mathbf{W}^{(i,j)}_R \right] \hat{\Omega}_p \left[ \mathbf{W}^{(i,j)}_1 \mathbf{W}^{(i,j)}_2 \ldots \mathbf{W}^{(i,j)}_R \right]' (32),
\]

where \( \hat{\Omega}_p \) is from minus the inverse of the Hessian of the likelihood in the usual fashion.

Coevolution models, i.e., models where (network) connectivity, \( W \), is some function of (behavior) \( y \), present larger challenges. Our simple stratagem for a first cut is to apply the poor man’s exogeneity: we lag temporally the \( y \) in this function explaining \( W \) and assume the conditions required for that identification-strategy to suffice hold adequately. As noted above, this does not address the problem of true or effective simultaneity, which seems likely at least in C&IPE contexts. Accordingly, 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 \( \hat{W} \) to insert in the likelihood, the 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 \( \hat{W} \). 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.\(^{41}\) 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: \(F \beta\). The simultaneity 
asymptotic biases (inconsistencies) of the naïve S-OLS estimator, which simply inserts the 
multiple spatial lags into least-squares regression, are also analogous to (8) as follows:

\[
\begin{align*}
\text{Let } y &= Z \rho + \varepsilon \quad \text{where } Z = \begin{bmatrix} W_{1y} & W_{2y} & \cdots & W_{ny} \end{bmatrix} \text{ and } \rho = \begin{bmatrix} \rho_1 & \rho_2 \end{bmatrix}. \\
\text{Then: } \text{plim } \hat{\rho} &= \rho + \text{plim } \left[ \frac{Z'Z}{n} \right]^{-1} \frac{Z'\varepsilon}{n} \quad \text{that is:} \\
\text{plim } \hat{\rho} &= \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} + \begin{bmatrix} V(W_{1y}) & -C(W_{1y},W_{1y}) \\ \vdots & \vdots \\ V(W_{ny}) & -C(W_{ny},W_{ny}) \end{bmatrix} \begin{bmatrix} V(W_{1y}) \times V(W_{1y}) \times C(W_{1y},W_{1y}) \\ \vdots \\ V(W_{ny}) \times V(W_{ny}) \times C(W_{ny},W_{ny}) \end{bmatrix}^{-1} \begin{bmatrix} C(\varepsilon, W_{1y}) \\ \vdots \\ C(\varepsilon, W_{ny}) \end{bmatrix} \\
\text{that is: } \text{plim } \hat{\rho} &= \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} + \begin{bmatrix} \Psi_1 \Gamma_1 \Psi_2 - \Psi_2 \Gamma_2 \\ \vdots \\ \Psi_n \Gamma_n - \Psi_n \Gamma_n \end{bmatrix}, \text{ where } \Psi = \text{plim } \left[ \frac{Z'Z}{n} \right] \text{ and } \Gamma = \text{plim } \left[ \frac{Z'\varepsilon}{n} \right]. 
\end{align*}
\]

The intuitions remain as before: simultaneity biases generally increase in \(\rho\), concentrating in 
the spatial lags that covary most with the residual, and inducing biases in generally opposite 
directions for other spatial-lag and covariates’ coefficient-estimates.

To demonstrate that an estimator is inconsistent, however, does not demonstrate that 
these asymptotic biases 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 
bias, inefficiency, and standard-error inaccuracy. With the analytical results in (33), Table 
1’s simulation results are easily interpreted. The covariance of the \textit{queen} spatial-lag (all eight 
adjacent squares on a grid) and \(\varepsilon\) is about half that of the \textit{rook} lag (only the four horizontally 
and vertically adjacent) and \(\varepsilon\). With row-standardization, eight ties, and fixed \(\rho\), the strength

\(^{41}\) Neither do we as-yet have adequately functioning proof of concept, in fact.
of the interdependence/endogeneity is more diluted in queen-lag $W_2y$. Consequently, $\hat{\rho}_1$ on rook-lag $W_1y$ is overestimated, and $\hat{\rho}_2$ on queen-lag $W_2y$ and $\hat{\beta}_0$ on $x_0$ (which is especially correlated with $W_2y$)\(^{42}\) are (badly) underestimated. The S-ML estimator also dominates impressively in efficiency and standard-error accuracy, especially for those three coefficients.

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

<table>
<thead>
<tr>
<th>ESTIMATOR</th>
<th>RESULT</th>
<th>$\beta_0=1$</th>
<th>$\beta_1=1$</th>
<th>$\rho_1=.3$</th>
<th>$\rho_2=.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-OLS</td>
<td>Average Estimate</td>
<td>.38 / .24</td>
<td>.96 / .97</td>
<td>.47 / .47</td>
<td>.27 / .29</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.33 / .24</td>
<td>.07 / .05</td>
<td>.19 / .15</td>
<td>.21 / .16</td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.71 / .80</td>
<td>.08 / .06</td>
<td>.25 / .22</td>
<td>.21 / .16</td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.37 / .28</td>
<td>.06 / .05</td>
<td>.15 / .11</td>
<td>.17 / .13</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>.92 / .87</td>
<td>1.06 / 1.02</td>
<td>1.29 / 1.35</td>
<td>1.21 / 1.27</td>
</tr>
<tr>
<td>S-ML</td>
<td>Average Estimate</td>
<td>1.09 / 1.05</td>
<td>1.00 / 1.00</td>
<td>.31 / .31</td>
<td>.27 / .28</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.33 / .24</td>
<td>.07 / .05</td>
<td>.12 / .09</td>
<td>.14 / .11</td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.34 / .24</td>
<td>.07 / .05</td>
<td>.12 / .09</td>
<td>.14 / .11</td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.31 / .23</td>
<td>.06 / .05</td>
<td>.12 / .09</td>
<td>.14 / .11</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>1.05 / 1.05</td>
<td>1.03 / 1.00</td>
<td>.98 / .91</td>
<td>.98 / 1.01</td>
</tr>
</tbody>
</table>

Monte Carlo (1000 Trials) Results for $y=\rho_1W_1y+\rho_2W_2y+\beta_0x_0+\varepsilon$, with $W_1=$rook adjacency, $W_2=$queen adjacency (row normalized); $\beta_0=\beta_1=1$, $\rho_0=\rho_1=.3$; and $N=225/450$.

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

<table>
<thead>
<tr>
<th>ESTIMATOR</th>
<th>RESULT</th>
<th>$\phi=.3$</th>
<th>$\beta=1$</th>
<th>$\rho=.3$</th>
<th>$\gamma=.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>Average Estimate</td>
<td>.36</td>
<td>1.04</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.07</td>
<td>.05</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.09</td>
<td>.06</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.03</td>
<td>.05</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>2.07</td>
<td>.97</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>S-OLS</td>
<td>Average Estimate</td>
<td>.28</td>
<td>.99</td>
<td>.41</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.03</td>
<td>.05</td>
<td>.06</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.04</td>
<td>.05</td>
<td>.129</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.03</td>
<td>.04</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>1.02</td>
<td>1.05</td>
<td>1.19</td>
<td>1.01</td>
</tr>
<tr>
<td>S-ML</td>
<td>Average Estimate</td>
<td>.29</td>
<td>1.00</td>
<td>.31</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.03</td>
<td>.05</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.03</td>
<td>.05</td>
<td>.05</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.03</td>
<td>.04</td>
<td>.04</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>1.02</td>
<td>1.04</td>
<td>1.11</td>
<td>1.10</td>
</tr>
</tbody>
</table>

Monte Carlo (1000 Trials) Results for $y_1=\rho Wy_1+\gamma Ly_1+\phi y_{1-1}+\beta x_1+\varepsilon$, with $W=48$ contiguous US-state adjacency pattern (row-stdzd); $\rho=.3$, $\gamma=.3$, $\phi=.3$, $\beta=1$; and $N=48$, $T=10$.

Table 2 similarly evaluates the blind OLS, naïve S-OLS, and S-ML estimators for our m-STAR coevolution model. S-ML again outperforms the inconsistent OLS alternatives. Notice that, with $x$ 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 $\beta$ and concentrate at a noticeable 20% in $\phi$. Notice also that with endogenous $L_y$ being time-lagged as estimated and in truth, and $L$ being $|y_i-y_j|$ and so not terribly linearly related to $y_i$, the simultaneity biases of S-OLS concentrate in $\rho$ at a sizable +33%, whereas $x$ being drawn

\(^{42}\) The simulations drew $\varepsilon$ from $N(1,1)$, making the nonzero aspect of $ZWy$ concentrate in the constant, $x_0$. 

Page 22 of 36
independently, little induced bias appears in $\hat{\beta}$. The efficiency (RMSE) and standard-error accuracy gains are more uniform and obvious.

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 model of 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. 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.

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

---

43 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.
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 evidences some spatial clustering on geographic, cultural, or 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 apparently spatially clustered 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$)

<table>
<thead>
<tr>
<th>Country</th>
<th>AUS</th>
<th>AUT</th>
<th>BEL</th>
<th>CAN</th>
<th>DEN</th>
<th>FIN</th>
<th>FRA</th>
<th>DEU</th>
<th>GRE</th>
<th>IRE</th>
<th>ITA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment services &amp; admin</td>
<td>28.85</td>
<td>33.06</td>
<td>41.35</td>
<td>38.44</td>
<td>20.48</td>
<td>22.47</td>
<td>25.98</td>
<td>46.59</td>
<td>10.77</td>
<td>32.36</td>
<td>0</td>
</tr>
<tr>
<td>Labor-market training</td>
<td>10.05</td>
<td>41.58</td>
<td>43.99</td>
<td>46.18</td>
<td>117.32</td>
<td>47</td>
<td>49.74</td>
<td>71.61</td>
<td>7.84</td>
<td>43.26</td>
<td>7.14</td>
</tr>
<tr>
<td>Youth measures</td>
<td>8.36</td>
<td>4.63</td>
<td>3.41</td>
<td>3.92</td>
<td>24.57</td>
<td>17.65</td>
<td>42.47</td>
<td>12.83</td>
<td>6.8</td>
<td>37.11</td>
<td>31.47</td>
</tr>
<tr>
<td>Subsidized employment</td>
<td>21.51</td>
<td>18.62</td>
<td>14.9</td>
<td>7.07</td>
<td>69.26</td>
<td>87.79</td>
<td>49.75</td>
<td>65.7</td>
<td>8.28</td>
<td>87.35</td>
<td>24.68</td>
</tr>
<tr>
<td>Measures for disabled</td>
<td>8.08</td>
<td>13.51</td>
<td>27.94</td>
<td>2.7</td>
<td>55.98</td>
<td>19.12</td>
<td>12.56</td>
<td>46.95</td>
<td>1.28</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Total ALMP</td>
<td>76.87</td>
<td>111.41</td>
<td>263.61</td>
<td>98.3</td>
<td>287.2</td>
<td>214.81</td>
<td>180.51</td>
<td>243.72</td>
<td>34.97</td>
<td>213.08</td>
<td>88.97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country</th>
<th>JPN</th>
<th>NTH</th>
<th>NWZ</th>
<th>NOR</th>
<th>PRT</th>
<th>ESP</th>
<th>SWE</th>
<th>CHE</th>
<th>GBR</th>
<th>USA</th>
<th>OECD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment services &amp; admin</td>
<td>47.65</td>
<td>24.48</td>
<td>17.46</td>
<td>33.36</td>
<td>14.93</td>
<td>12.97</td>
<td>44.98</td>
<td>20.81</td>
<td>36.4</td>
<td>12.1</td>
<td>26.76</td>
</tr>
<tr>
<td>Labor-market training</td>
<td>7.48</td>
<td>48.72</td>
<td>39.62</td>
<td>36.28</td>
<td>26.47</td>
<td>19.23</td>
<td>101.86</td>
<td>17.01</td>
<td>15.5</td>
<td>13.48</td>
<td>41.12</td>
</tr>
<tr>
<td>Youth measures</td>
<td>0.24</td>
<td>11.01</td>
<td>10.44</td>
<td>11.53</td>
<td>26.48</td>
<td>7.73</td>
<td>14.32</td>
<td>0.55</td>
<td>28.34</td>
<td>6.68</td>
<td>14.84</td>
</tr>
<tr>
<td>Subsidized employment</td>
<td>12.73</td>
<td>16.69</td>
<td>34.73</td>
<td>23.87</td>
<td>26.47</td>
<td>14.77</td>
<td>45.7</td>
<td>91.75</td>
<td>20.05</td>
<td>11.68</td>
<td>2.24</td>
</tr>
<tr>
<td>Measures for disabled</td>
<td>1.53</td>
<td>110.57</td>
<td>37.91</td>
<td>104.59</td>
<td>4.95</td>
<td>2.64</td>
<td>107.98</td>
<td>33.79</td>
<td>4.61</td>
<td>9.22</td>
<td>29.29</td>
</tr>
<tr>
<td>Total ALMP</td>
<td>69.62</td>
<td>211.47</td>
<td>110.15</td>
<td>209.63</td>
<td>87.59</td>
<td>88.28</td>
<td>360.88</td>
<td>96.52</td>
<td>43.72</td>
<td>154.3</td>
<td></td>
</tr>
</tbody>
</table>

**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. The model, in matrix notation, is

\[
y = \sum_{r=1}^{R} \rho_r W_r y + \phi M y + \gamma L y + X \beta + \epsilon
\]

(34),

where \( y \), ALM expenditures, is an \( NT \times 1 \) vector of cross sections stacked by periods as described previously. \( \rho_r \) is the \( r^{th} \) spatial autoregressive coefficient, and \( W_r \) is an \( NT \times NT \) block-diagonal spatial-weighting matrix. Each \( W_r \) contains a unique set of elements \( w_{ij} \) that reflect a

44 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.
particular type of interdependence (e.g., geographic proximity, EU co-membership, and economic interdependence). In the other new term, $\gamma$, is the coevolutionary-dependence parameter, and $L$ is an $NT \times NT$ “policy-distance” matrix with $|y_{i,t-1} - y_{j,t-1}|$ in cells $(it,jt)$. The addition of $\gamma Ly$, 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 policies. In spatial-econometric terms, it is the endogenous determinant of the strength of interdependence. The rest is as before: $Wy$ reflects the exogenous interdependence of units’ policies; $My$ is the first-order temporal lag, with $\phi$ its coefficient; $X\beta$ are the exogenous non-spatial components; and $\varepsilon$ are assumed i.i.d. disturbances.

The presence in (34) of $L$, 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 $X$ or $\delta$ must be calculated recursively. To start, rewrite (34) as $t$ cross-sections:

$$y_t = Wy_t + \phi y_{t-1} + \gamma \left[ abs \left( \Pi \left[ y_{t-1} \otimes I_k \right] \right) \right] y_t + X\beta + \varepsilon_t \tag{35}$$

$y_t$, $Wy_t$, and $X_t$ are $N \times 1$, $N \times N$, and $N \times k$ matrices, and $\Pi$ is an $N \times N^e$ matrix produced by horizontally concatenating $N$ separate $N \times N$ matrices. The $i^{th}$ $N \times N$ matrix in $\Pi$ 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, $\Pi$ looks like this:

$$\Pi = \begin{bmatrix}
0 & 0 & 0 & -1 & 1 & 0 & -1 & 0 & 1 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & -1 & 0 & 1 & -1 & 0 & 0 \\
\end{bmatrix} \tag{36}$$

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

$$y_t = \left[ I_n - Wy_t - \gamma \left[ abs \left( y_{t-1} \otimes I_k \right) \right] \right]^{-1} (\phi y_{t-1} + X\beta + \varepsilon_t) \tag{37}$$

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, $Wy$, involve four different weights matrices ($R=4$). $W_1$ 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. $W_2$ 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. $W_4$ has weights that reflect the trade (imports+exports) shares between sample countries. $W_4$ is the policy-distance coevolution matrix with $|y_{it}-y_{jt}|$ in cells $(it,jt)$ as described above. For estimation, we row-standardize (as is common in spatial econometrics) all $W$ matrices, dividing each cell by its row’s sum.

We also control several domestic variables. We include macroeconomic performance in the form of real GDP-growth and unemployment rates. Wealthier political-economies may demand more public goods and services from their governments, suggesting a positive coefficient 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, may suggest a negative 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 should increase the influence of organized labor, so we expect it to associate 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 spending in particular. 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 percents of cabinet seats held by left-wing and by Christian Democratic parties, and the percent of general-election 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

---

45 More specifically, $w_{ij}(t)$ equals the sum of exports $i$ to $j$ and $j$ to $i$ and of imports 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.

46 Row normalization is standard practice in spatial econometrics and mathematically/analytically convenient, even though it is not necessarily substantively neutral (see, e.g., Plümper & Neumayer 2008).
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

<table>
<thead>
<tr>
<th>DEP. VAR.</th>
<th>Total ALM</th>
<th>LMT</th>
<th>SEMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDEP. VAR.</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Constant</td>
<td>-195.89*</td>
<td>-150.369</td>
<td>51.487</td>
</tr>
<tr>
<td>Temporal Lag</td>
<td>0.875***</td>
<td>0.836***</td>
<td>0.801***</td>
</tr>
<tr>
<td>Real GDP Growth Rate</td>
<td>1.365***</td>
<td>0.449</td>
<td>-1.100</td>
</tr>
<tr>
<td>Stnd. Unemployment Rate</td>
<td>-0.070</td>
<td>0.552</td>
<td>0.125</td>
</tr>
<tr>
<td>Union Density</td>
<td>0.888***</td>
<td>0.711***</td>
<td>0.527***</td>
</tr>
<tr>
<td>Deindustrialization</td>
<td>1.259</td>
<td>1.209</td>
<td>0.499</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>-0.522***</td>
<td>-0.183</td>
<td>-0.192</td>
</tr>
<tr>
<td>Working-Age Population</td>
<td>0.946</td>
<td>0.216</td>
<td>0.108</td>
</tr>
<tr>
<td>Left Cabinet-Seats</td>
<td>-0.024</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Christian Dem. Cab-Seats</td>
<td>-0.160*</td>
<td>-0.102</td>
<td>-0.085</td>
</tr>
<tr>
<td>Left-Libertarian Vote</td>
<td>-0.285</td>
<td>-0.549</td>
<td>-0.728</td>
</tr>
</tbody>
</table>

SPATIAL WEIGHTS:

Borders | -0.112*** | -0.098** | -0.185*** | -0.188*** | 0.058 | 0.054 |
| EU Membership | -0.071*** | -0.060*** | -0.079 | -0.183*** | -0.058 | -0.097 |
| Trade Shares | 0.386*** | 0.239*** | 0.316*** | 0.180 | -0.017 | -0.097 |
| Policy Distance | -0.053 | -0.189*** | -0.004 | -0.298*** | 0.049 | -0.205*** |

TIME DUMMIES:  

<table>
<thead>
<tr>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
</table>


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 models 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-period 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 in obtaining good estimates of the strength of interdependence, i.e., of $p$, 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, country and year dummies serve as a powerfully conservative means of accounting common outside shocks or spatially correlated domestic factors. Failure to account for such alternatives will bias spatial-lag coefficient-estimates, usually positively.\footnote{The claim of conservatism in this strategy arises from a kind of spatial-Nickell/Hurwicz bias that we have found to operate with time fixed-effects and spatial lags (see Franzese & Hays 2004, 2006, 2007, 2008).}

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, oddly indicating strong procyclicality to total ALM (though not LMT or SEMP components), and to labor-market structure, deindustrialization and especially union density. The spatial models suggest that the effects attributed by the non-spatial models to domestic growth and trade exposure seem instead to reflect spatial diffusion of responses to global conditions. Interestingly, all three estimation techniques find sizable differences in sources of LMT or 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 sensibly 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: LMT associates 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 $p_r$ estimates being significant at conventional levels and the six joint tests of the four $p_r$ estimates per spatial model all overwhelmingly rejecting null hypotheses of zero coefficients, i.e., of the nonspatial model.\footnote{Likelihood-ratio tests of the models and information criteria also strongly favor the spatial models.} In particular, total ALM spending seems strongly spatially interdependent on all four dimensions of proximity, SEMP much less so, and LMT intermediately. Consequently, coefficient estimates in non-spatial models will almost certainly suffer omitted-variable biases (of magnitudes in the same relative ordering).

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 \textit{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: in other (network-analytic) terms, \textit{homophily}.

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 \textit{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: \[ \mathbf{I}_n - \mathbf{W} - \gamma \left[ \text{abs} \left( \Pi \left[ \mathbf{y}_{t-1} \otimes \mathbf{I}_n \right] \right) \right]^{-1}, \] 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, \( \mathbf{y}_{t-1} \), 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 $W$ that we will assume to maintain in the long-run or obtain over the period in question. With endogenously time-varying $W$, like our policy-distance matrix $W_t$, 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>
<thead>
<tr>
<th>Table 5: Effects of a Common Counterfactual ($1$) Shock to ALM Spending</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Shock Steady-State ALM-Spending ($$$)</strong></td>
</tr>
<tr>
<td>Australia</td>
</tr>
<tr>
<td>Austria</td>
</tr>
<tr>
<td>Belgium</td>
</tr>
<tr>
<td>Canada</td>
</tr>
<tr>
<td>Denmark</td>
</tr>
<tr>
<td>Finland</td>
</tr>
<tr>
<td>France</td>
</tr>
<tr>
<td>Germany</td>
</tr>
<tr>
<td>Greece</td>
</tr>
<tr>
<td>Ireland</td>
</tr>
<tr>
<td>Italy</td>
</tr>
<tr>
<td>Japan</td>
</tr>
<tr>
<td>Netherlands</td>
</tr>
<tr>
<td>New Zealand</td>
</tr>
<tr>
<td>Norway</td>
</tr>
<tr>
<td>Portugal</td>
</tr>
<tr>
<td>Spain</td>
</tr>
<tr>
<td>Sweden</td>
</tr>
<tr>
<td>Switzerland</td>
</tr>
<tr>
<td>United Kingdom</td>
</tr>
<tr>
<td>United States</td>
</tr>
</tbody>
</table>

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 $\delta$) 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.
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 $10 permanent positive shock to Germany’s ALM spending ($\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 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.\(^49\)

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.\(^50\) EU member-countries are circles; other countries are squares. Red arrows represent negative interdependence, or implicit

\(^{49}\) Tables for the 1991 and 2001 values are too large to show effectively but are available on request.

\(^{50}\) The web appendix shows the analogous graphs using the 1981 and 2001 covariate values.
(net) network ties. Blue arrows indicate positive implicit (net) network-ties. Arrow 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.

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

<table>
<thead>
<tr>
<th></th>
<th>AUS</th>
<th>CAN</th>
<th>FIN</th>
<th>FRA</th>
<th>NTH</th>
<th>NWZ</th>
<th>ESP</th>
<th>SWE</th>
<th>GBR</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUS</td>
<td>0.000</td>
<td>0.022</td>
<td>-0.371***</td>
<td>0.160***</td>
<td>-0.182*</td>
<td>-0.905***</td>
<td>-0.025</td>
<td>-0.481***</td>
<td>0.179</td>
<td>1.123**</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.095)</td>
<td>(0.140)</td>
<td>(0.075)</td>
<td>(0.096)</td>
<td>(0.326)</td>
<td>(0.023)</td>
<td>(0.195)</td>
<td>(0.200)</td>
<td>(0.566)</td>
</tr>
<tr>
<td>CAN</td>
<td>-0.177**</td>
<td>0.000</td>
<td>-0.351***</td>
<td>-0.164**</td>
<td>-0.129**</td>
<td>-0.997***</td>
<td>-0.120***</td>
<td>-0.553***</td>
<td>0.020</td>
<td>1.092</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.000)</td>
<td>(0.126)</td>
<td>(0.075)</td>
<td>(0.055)</td>
<td>(0.036)</td>
<td>(0.046)</td>
<td>(0.200)</td>
<td>(0.064)</td>
<td>(0.787)</td>
</tr>
<tr>
<td>FIN</td>
<td>-0.283**</td>
<td>-0.159**</td>
<td>0.000</td>
<td>-0.072</td>
<td>0.080</td>
<td>-0.140***</td>
<td>-0.227***</td>
<td>-0.246</td>
<td>0.493</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.075)</td>
<td>(0.000)</td>
<td>(0.165)</td>
<td>(0.102)</td>
<td>(0.051)</td>
<td>(0.101)</td>
<td>(0.317)</td>
<td>(0.308)</td>
<td>(0.180)</td>
</tr>
<tr>
<td>FRA</td>
<td>0.047**</td>
<td>-0.071</td>
<td>-0.347***</td>
<td>0.000</td>
<td>-0.033</td>
<td>-0.207***</td>
<td>-0.275*</td>
<td>-0.418**</td>
<td>-0.345</td>
<td>0.565*</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.062)</td>
<td>(0.142)</td>
<td>(0.000)</td>
<td>(0.250)</td>
<td>(0.076)</td>
<td>(0.145)</td>
<td>(0.204)</td>
<td>(0.222)</td>
<td>(0.306)</td>
</tr>
<tr>
<td>NTH</td>
<td>-0.293**</td>
<td>-0.082</td>
<td>-0.134*</td>
<td>0.051</td>
<td>0.000</td>
<td>-0.039**</td>
<td>-0.175*</td>
<td>-0.238</td>
<td>-0.531*</td>
<td>0.357</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.056)</td>
<td>(0.070)</td>
<td>(0.313)</td>
<td>(0.000)</td>
<td>(0.016)</td>
<td>(0.099)</td>
<td>(0.149)</td>
<td>(0.290)</td>
<td>(0.280)</td>
</tr>
<tr>
<td>NWZ</td>
<td>-0.463</td>
<td>0.001</td>
<td>-0.244***</td>
<td>-0.226*</td>
<td>-0.005</td>
<td>0.000</td>
<td>-0.214***</td>
<td>-0.439***</td>
<td>0.575**</td>
<td>0.529</td>
</tr>
<tr>
<td></td>
<td>(0.328)</td>
<td>(0.062)</td>
<td>(0.089)</td>
<td>(0.116)</td>
<td>(0.028)</td>
<td>(0.000)</td>
<td>(0.082)</td>
<td>(0.163)</td>
<td>(0.272)</td>
<td>(0.343)</td>
</tr>
<tr>
<td>ESP</td>
<td>-0.038</td>
<td>-0.055</td>
<td>-0.381***</td>
<td>-0.220</td>
<td>-0.053</td>
<td>-0.190***</td>
<td>0.000</td>
<td>-0.502***</td>
<td>0.240</td>
<td>0.720**</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.047)</td>
<td>(0.146)</td>
<td>(0.308)</td>
<td>(0.124)</td>
<td>(0.069)</td>
<td>(0.099)</td>
<td>(0.214)</td>
<td>(0.215)</td>
<td>(0.355)</td>
</tr>
<tr>
<td>SWE</td>
<td>-0.243**</td>
<td>-0.148*</td>
<td>-0.638**</td>
<td>0.033</td>
<td>0.113</td>
<td>-0.169***</td>
<td>-0.190*</td>
<td>0.000</td>
<td>0.560</td>
<td>0.203</td>
</tr>
<tr>
<td></td>
<td>(0.106)</td>
<td>(0.083)</td>
<td>(0.299)</td>
<td>(0.188)</td>
<td>(0.141)</td>
<td>(0.062)</td>
<td>(0.101)</td>
<td>(0.000)</td>
<td>(0.354)</td>
<td>(0.231)</td>
</tr>
<tr>
<td>GBR</td>
<td>-0.215*</td>
<td>0.057</td>
<td>-0.175*</td>
<td>-0.619***</td>
<td>-0.405**</td>
<td>0.034*</td>
<td>-0.124</td>
<td>-0.270</td>
<td>0.000</td>
<td>0.633</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.080)</td>
<td>(0.104)</td>
<td>(0.218)</td>
<td>(0.186)</td>
<td>(0.021)</td>
<td>(0.094)</td>
<td>(0.198)</td>
<td>(0.000)</td>
<td>(0.387)</td>
</tr>
<tr>
<td>USA</td>
<td>0.008</td>
<td>0.284</td>
<td>-0.382***</td>
<td>0.108</td>
<td>-0.069</td>
<td>-0.146**</td>
<td>0.040</td>
<td>-0.54***</td>
<td>0.217</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.455)</td>
<td>(0.143)</td>
<td>(0.111)</td>
<td>(0.106)</td>
<td>(0.062)</td>
<td>(0.042)</td>
<td>(0.214)</td>
<td>(0.190)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

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

Figure 4: 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


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.


