Statistical Methodology 7 (2010) 406–428

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

A spatial model incorporating dynamic, endogenous network interdependence: A political science application[☆]

Jude C. Hays[a], Aya Kachi[a], Robert J. Franzese Jr.[b,*]

[a] Department of Political Science, University of Illinois, Urbana-Champaign, United States
[b] Department of Political Science, The University of Michigan, Ann Arbor, United States

ARTICLE INFO

Article history:
Received 7 May 2009
Received in revised form
30 November 2009
Accepted 30 November 2009

Keywords:
Interdependence
Network dependence
Spatial lag
Spatial econometrics

ABSTRACT

We propose applying 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 (*selection*). We 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: common exposure, contagion, and *selection* (e.g., *homophily*). We illustrate this approach to dynamic, endogenous interdependence – which parallels models of network–behavior co-evolution in the longitudinal networks literature – with an empirical application that aims to disentangle the roles of economic interdependence, correlated external and internal stimuli, and EU membership in shaping labor market policies in developed democracies in recent years.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Social scientific interest in and application of spatial modeling has burgeoned lately.[1] This is a welcome development because the dependence of outcomes in some units on outcomes in others (or the behavior of some actors on the behavior of others), i.e., *spatial interdependence*, is substantively

☆ An Appendix with related material is available at http://www.umich.edu/~franzese.

* Corresponding address: University of Michigan, Political Science, 426 Thompson Street, 48104 Ann Arbor, MI, United States. Tel.: +1 734 936 1850; fax: +1 734 764 3341.

E-mail addresses: jchays@uiuc.edu (J.C. Hays), akachi2@illinois.edu (A. Kachi), franzese@umich.edu (R.J. Franzese Jr.).

[1] The Appendix provides, *inter alia*, a large topically organized reference list for spatial econometric work in social science, emphasizing political science, and a complete reference list.

1572-3127/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.stamet.2009.11.005

ubiquitous and theoretically central across the social sciences. At the same time, the study of *network interdependence* has emerged as a central topic of research in a number of social scientific disciplines.

Despite the obvious conceptual, theoretical, and substantive overlap between spatial and network interdependence, spatial econometric and network analytic research have developed largely in isolation.[2] This is partly because spatial econometricians and network analysts tend to ask slightly different questions in their research and to have correspondingly differing objectives. The former tend to analyze the effects of spatial interdependence on outcomes or behaviors (i.e., *contagion*), taking the pattern of connectivity as given, while the latter emphasize analysis of network formation (i.e., *selection*), taking actors' attributes and behaviors as given. Furthermore, spatial econometricians tend primarily to conceive such *spatial/network effects* as the effects of alters' actions on ego's actions via their spatial/network connections, whereas network analysts tend to emphasize the effects on actions of network structure and ego's position in that network structure. Although both explaining networks and understanding how network interdependence influences outcomes generally interest social scientists, and although many have called for synthesis in empirical analyses of *contagion* and *selection*, as yet very few models can account for contagion and selection simultaneously.

From the perspective of social network analysis, Snijders and colleagues' co-evolution model (as reviewed below[3]) represents the furthest advance toward combining contagion and selection. For many other kinds of social science applications, though, the model has some significant limitations, particularly for research in *comparative and international political economy* (C & IPE), an area highlighted by our illustration. To begin to reduce some of these limitations, we suggest the *multiparametric spatiotemporal autoregressive* (m-STAR) model as a spatial econometric analogue to the network co-evolution model. Section 2 next offers introduction to the specification, estimation, and interpretation of baseline *spatiotemporal autoregressive* (STAR) models. Ensuing sections will then show how to apply the m-STAR extension of STAR models as a means to estimate simultaneously the patterns of interdependence (i.e., *network selection*) and the behavioral responses of units to other units' behaviors (i.e., *contagion*), including the possibility of endogenizing the pattern of interdependence (i.e., the network) dynamically to the outcome (i.e., the possibility of *network–behavior co-evolution*).

2. Spatiotemporal models of interdependence: Specification, estimation, interpretation

Analysis of time-series cross-section (TSCS) data, which typically manifest both temporal and spatial dependence importantly, is common in the social sciences. In C & IPE, for instance, interdependence often arises strongly, and substantively and theoretically centrally, as it does in the contexts of policy diffusion and of globalization, tax competition, and policy autonomy, for example. To estimate effects and draw inferences validly in such contexts, analysts generally do best to specify both temporal and spatial dependence directly in their empirical models by using spatial and temporal lags.[4] A baseline such *spatiotemporal autoregressive* (STAR) model could include, for instance, a first-order temporal lag and a contemporaneous spatial lag:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \phi\mathbf{M}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{1}$$

The dependent variable, $\mathbf{y}$, is an $NT \times 1$ vector of cross-sections stacked by periods (i.e., the N first-period observations through the N in period T). Next, ρ is the spatial autoregressive coefficient, giving the strength of interdependence operating along the pattern given by $\mathbf{W}_{NT}$, the $NT \times NT$ block diagonal spatial weighting matrix. Each of $N \times N$ weights matrices, $\mathbf{W}_{N,t}$, on the block diagonal of $\mathbf{W}_{NT}$ have elements $w_{ij,t}$ reflecting the relative connectivity from unit j to i in period t. Thus, for each y_{it}, the *spatial lag*, $\mathbf{Wy}$, gives a weighted sum of the y_{jt}, with weights $w_{ij,t}$: a direct and straightforward

[2] The same may be said, to a perhaps somewhat lesser extent, of spatial statistics and network analysis.

[3] The Appendix also provides further description of these network analytic co-evolution models.

[4] Some scholars use spatial methods to purge their data of spatial association; others use spatial methods to study these associations. We are more interested in the substantive implications of spatial and network patterns, and so adopt the latter approach.

reflection of the dependence of each unit i's outcome on other units' outcomes. (Note that $\mathbf{W}$ here is data, pre-specified and given.) $\mathbf{M}$ is an $NT \times NT$ matrix with ones on the lower first-minor diagonal, i.e., at coordinates $(N+1, 1), (N+2, 2), \ldots, (NT, NT-1)$, and zeros elsewhere. $\mathbf{My}$ is thus a typical first-order temporal lag, with ϕ its coefficient. $\mathbf{X}$ contains NT observations on k independent variables; $\boldsymbol{\beta}$ is its $k \times 1$ vector of coefficients; and $\boldsymbol{\varepsilon}$ is an $NT \times 1$ vector of *i.i.d.* stochastic components.[5]

The simultaneity of the spatial lag regressor, $\mathbf{Wy}$, would clearly bias OLS estimation of spatial autoregressive models (spatial OLS or S-OLS). Consistent and asymptotically efficient estimation by maximum likelihood is reasonably straightforward though.[6] The conditional likelihood for the STAR model [13–15], which assumes the first observations are non-stochastic, is a simple extension of the standard spatial autoregressive (or SAR) model likelihood. To see this, start by rewriting the SAR model with the stochastic component on the left:

$$\mathbf{y} = \rho\mathbf{Wy} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \Rightarrow \boldsymbol{\varepsilon} = (\mathbf{I} - \rho\mathbf{W})\,\mathbf{y} - \mathbf{X}\boldsymbol{\beta} \equiv \mathbf{Ay} - \mathbf{X}\boldsymbol{\beta}, \tag{2}$$

where $\mathbf{My}$, first-order time lag of $\mathbf{y}$, is now the first column of $\mathbf{X}$ and ϕ the first row of $\boldsymbol{\beta}$.[7] Assuming *i.i.d.* normality, the likelihood for $\boldsymbol{\varepsilon}$ is just the typical linear normal one:

$$L(\boldsymbol{\varepsilon}) = \left(1/\sigma^2 2\pi\right)^{\frac{NT}{2}} \exp\left(-\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/2\sigma^2\right), \tag{3}$$

which, applying the change of variables theorem, produces this likelihood in terms of $\mathbf{y}$:

$$L(\mathbf{y}) = |\mathbf{A}| \left(1/2\pi\sigma^2\right)^{\frac{NT}{2}} \exp\left(\left(-1/2\sigma^2\right)(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})\right). \tag{4}$$

This resembles the CLNRM likelihood, except that the transformation from $\boldsymbol{\varepsilon}$ to $\mathbf{y}$ is not by the usual factor, 1, but by $|\mathbf{A}| = |\mathbf{I} - \rho\mathbf{W}|$.[8] Written in $(N \times 1)$ vector notation, this STAR model conditional likelihood is conveniently separable into parts, like so:

$$\mathrm{Log} f_{\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{y}_2 | \mathbf{y}_1} = -\frac{1}{2} N\,(T-1) \log\left(2\pi\sigma^2\right) + (T-1) \log\left|\mathbf{A}_{N,t}\right| - \frac{1}{2\sigma^2} \sum_{t=2}^{T} \boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t \tag{5}$$

where

$$\mathbf{A}_{N,t} = (\mathbf{I}_N - \rho\mathbf{W}_{N,t}) \quad \text{and} \quad \boldsymbol{\varepsilon}_t = \mathbf{y}_t - \phi\mathbf{y}_{t-1} - \rho\mathbf{W}_{N,t}\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta}.$$

The unconditional (exact) likelihood function, which retains the first-period observations as non-predetermined, is more complex [15]:

$$\mathrm{Log} f_{\mathbf{y}_t, \ldots, \mathbf{y}_1} = -\frac{1}{2} NT \log\left(2\pi\sigma^2\right) + \frac{1}{2} \sum_{i=1}^{N} \log\left((1 - \rho\omega_i)^2 - \phi^2\right)$$

5 Higher order temporal dynamics would simply add further properly configured weights matrices, altering the implied spatiotemporal dynamics but not complicating estimation. Similarly, STAR models might also include temporally lagged spatial lag variables, like: $\mathbf{y}_t = \rho\mathbf{W}_{N,t}\mathbf{y}_t + \theta\mathbf{W}_{N,t-1}\mathbf{y}_{t-1} + \phi\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t$. This too would change the spatiotemporal dynamics implied by the model but not complicate its estimation. Likewise, spatial lag independent variables can be added (giving a spatiotemporal version of the so-called spatial Durbin model), again complicating the implied dynamics but not estimation. Other feasible complications ignored here include: non-rectangular data sets or missing data, non-zero off-diagonal blocks of $\mathbf{W}_{NT}$, and non-spherical error covariance structure (including spatial error components or spatial lag error models). All would complicate exposition greatly without commensurately enhancing expository illumination.

6 Franzese and Hays [17,18,21,22,24] explored these simultaneity biases of S-OLS and omitted-variable biases of OLS omitting the spatial lag analytically and by simulation, finding the latter are generally larger, often far larger, than the former. We also discussed consistent (and asymptotically efficient) estimation by S-2SLS (or S-GMM), which could proceed with $\mathbf{WX}$ as valid instrument for $\mathbf{Wy}$, given the already assumed exogeneity of $\mathbf{W}$ and $\mathbf{X}$. However, our simulation analyses have so far shown ML estimation to be nearly weakly dominant, so we describe only it next.

7 N.B., although $\mathbf{Wy}$ complicates the conditional likelihood in terms of $\mathbf{y}$ (see footnote 8), $\mathbf{My}$ does not.

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

$$+ (T-1)\log|\mathbf{A}_{N,t}| - \frac{1}{2\sigma^2}\sum_{t=2}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t - \frac{1}{2\sigma^2}\boldsymbol{\varepsilon}_1'\left(\left(\mathbf{A}_{N,1}-\phi\mathbf{I}_N\right)'\right)^{-1}$$
$$\times\left(\mathbf{A}_{N,1}'\mathbf{A}_{N,1}-\phi^2\mathbf{A}_{N,1}'\mathbf{A}_{N,1}^{-1}\left(\mathbf{A}_{N,1}'\mathbf{A}_{N,1}^{-1}\right)'\right)^{-1}\left(\mathbf{A}_{N,1}-\phi\mathbf{I}_N\right)^{-1}\boldsymbol{\varepsilon}_1 \tag{6}$$

where

$$\boldsymbol{\varepsilon}_1 = (\mathbf{I}_N-\phi\mathbf{I}_N)\,\mathbf{y}_1 - \rho\mathbf{W}_{N,1}\mathbf{y}_1 - \mathbf{x}_1'\boldsymbol{\beta}.$$

With large T, the first observation contributes little to the total likelihood, so scholars can safely use the simpler conditional likelihood for estimation.

One easy way to ease or even erase the simultaneity problem of S-OLS is to time lag the spatial lag (see, e.g., [5]). Insofar as time lagging the spatial lag renders it pre-determined – that is, as interdependence does not incur *instantaneously* (meaning: *within an observational period, as measured, given the model*) – OLS's bias disappears asymptotically ($T\to\infty$). Formally, the STAR model with time lagged spatial lag is

$$\mathbf{y}_t = \eta\mathbf{W}_{N,t-1}\mathbf{y}_{t-1} + \phi\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t. \tag{7}$$

Elhorst ([13]: 126–130) derives the unconditional log-likelihood for this model as

$$\mathrm{Log}f_{\mathbf{y}_t,\ldots,\mathbf{y}_1} = -\frac{1}{2}NT\log\left(2\pi\sigma^2\right) + \frac{1}{2}\sum_{i=1}^{N}\log\left(1-(\phi+\eta\omega_i)^2\right) - \frac{1}{2\sigma^2}\sum_{t=2}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t$$
$$-\frac{1}{2\sigma^2}\boldsymbol{\varepsilon}_1'\left(\left(\mathbf{I}_N-\mathbf{B}_{N,1}\right)'\right)^{-1}\left(\mathbf{I}_N-\mathbf{B}_{N,1}\mathbf{B}_{N,1}'\right)^{-1}\left(\mathbf{I}-\mathbf{B}_{N,1}\right)^{-1}\boldsymbol{\varepsilon}_1, \tag{8}$$

with $\boldsymbol{\varepsilon}_1 = \mathbf{y}_1 - (\phi+\eta\mathbf{W}_{N,1})\mathbf{y}_1 - \mathbf{X}_1\boldsymbol{\beta}$; $\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \eta\mathbf{W}_{N,t-1}\mathbf{y}_{t-1} - \phi\mathbf{y}_{t-1} - \mathbf{X}_t\boldsymbol{\beta}$; $\mathbf{B}_{N,1} = \phi\mathbf{I}_N + \eta\mathbf{W}_{N,1}$. Note that the second and last terms in (8) bias OLS estimation of the time lagged spatial lag STAR model, although this bias goes to zero as $T\to\infty$. Thus, (i) if T is large, (ii) if spatial interdependence processes do not operate within an observational period but only with a time lag (as measured and modeled), and (iii) if spatiotemporal dynamics are modeled well enough for these conditions not to become violated through measurement *cum* misspecification error, then simple OLS with time lagged spatial lag regressor(s) is a valid and effective estimation strategy. In many social scientific contexts, however, these conditions for valid OLS estimation greatly strain substantive credulity.

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

$$\left|\phi\left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}\right| < 1, \quad \text{or, equivalently, if } \begin{cases}|\phi| < 1-\rho\omega_{\max}, & \text{if } \rho\geq 0\\ |\phi| < 1-\rho\omega_{\min}, & \text{if } \rho < 0.\end{cases} \tag{9}$$

For instance, with positive temporal and spatial dependence and $\mathbf{W}$ row standardized, $\omega_{\max}=1$, and 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 and separable, involves more than just eyeing coefficient estimates. With spatiotemporal dynamics, as with solely temporal dynamics, coefficients on other regressors give only the pre-dynamic impulses from those variables to the outcome. That is, coefficients represent only the pre-interdependence feedback *impetus* (often inherently unobservable) to outcomes from other regressors. For fuller substantive interpretation of spatiotemporally dynamic models, one must calculate spatiotemporal multipliers (as explained next). These multipliers allow expression of estimated responses of the dependent variable across all units, accounting the spatiotemporal dynamics, to shocks to covariates or errors in any unit(s). They also afford estimation of the long run, steady state, or equilibrium[9] effects of permanent such shocks.[10] We apply the delta method to derive analytically the approximate asymptotic variance–covariance matrix (standard

[9] We use the terms *long run*, *steady state*, and *equilibrium* effects loosely, interchangeably, to refer to the estimated asymptotic level of outcomes **y** following a hypothetical *permanent* shock.

[10] Anselin [3], Franzese and Hays [17,18,21,22,24], LeSage and Pace [50] discuss these multipliers more fully.

errors) for these response path and long run effect estimates; these standard errors could also be simulated of course.

One calculates cumulative, long run steady state (LRSS) spatiotemporal effects most conveniently working with the STAR model in $(N \times 1)$ vector form:

$$\mathbf{y}_t = \rho \mathbf{W}_{N,t}\mathbf{y}_t + \phi \mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t. \tag{10}$$

Setting $\mathbf{y}_{t-1}$ equal to $\mathbf{y}_t$ and solving produces the LRSS level of $\mathbf{y}$, assuming stationarity and exogenous RHS terms, $\mathbf{X}$ and/or $\boldsymbol{\varepsilon}$, permanently fixed to their hypothetical levels[11]:

$$\begin{aligned}
\mathbf{y}_t &= \rho \mathbf{W}_{N,t}\mathbf{y}_t + \phi \mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t = \left(\rho \mathbf{W}_{N,t} + \phi \mathbf{I}\right)\mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \\
&= \left[\mathbf{I}_N - \rho \mathbf{W}_{N,t} + \phi \mathbf{I}_N\right]^{-1}(\mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t) \\
&= \begin{bmatrix} 1-\phi & -\rho w_{1,2} & \cdots & \cdots & -\rho w_{1,N} \\ -\rho w_{2,1} & 1-\phi & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 1-\phi & -\rho w_{(N-1),N} \\ -\rho w_{N,1} & \cdots & \cdots & -\rho w_{N,(N-1)} & 1-\phi \end{bmatrix}_t^{-1} (\mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t) \\
&\equiv \mathbf{S} \times (\mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t).^{12}
\end{aligned} \tag{11}$$

(Decomposing $\boldsymbol{\varepsilon}_t = \boldsymbol{\delta} + \boldsymbol{\upsilon}_t$ with $\boldsymbol{\delta}$ fixed and $\boldsymbol{\upsilon}_t$ stochastic is conceptually useful in calculating responses to counterfactual shocks to the outcomes in some unit(s), which could then be considered shocks to $\boldsymbol{\delta}$.) Again, the delta or parametric bootstrap methods can provide standard error estimates for these estimated LRSS responses.

The spatiotemporal response path of the $N \times 1$ vector of unit outcomes, $\mathbf{y}_t$, to hypothetical shocks to the exogenous right hand side terms, $\mathbf{X}$ and $\boldsymbol{\varepsilon}$, could also emerge by rearranging (10) to isolate $\mathbf{y}_t$ on the left hand side:

$$\mathbf{y}_t = \left[\mathbf{I}_N - \rho \mathbf{W}_{N,t}\right]^{-1}[\phi \mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t] = \mathbf{S}[\phi \mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t]. \tag{12}$$

This formula gives the response paths of all units $\{i\}$ to counterfactual shocks to $\mathbf{X}$ or $\boldsymbol{\varepsilon}$ in any unit(s) $\{j\}$, including shocks in $\{i\}$ itself/themselves, by setting $(\mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t)$ to the values reflecting that hypothetical in row(s) $\{j\}$. For the marginal (i.e., non-cumulative) spatiotemporal effects or response paths, and their standard errors, working in $NT \times NT$ matrix form is easier. Simply redefine $\mathbf{S}$ in (11) as $\mathbf{S} \equiv [\mathbf{I}_{NT} - \rho \mathbf{W} - \phi \mathbf{M}]^{-1}$ and follow the same steps as are given thereafter. We illustrate calculation of estimated spatiotemporal responses (with delta method standard errors) in presenting our empirical application below.[13]

3. The multiparametric spatiotemporal (m-STAR) model, with co-evolution

Given the crucial role of the quality of $\mathbf{W}$'s pre-specification to estimation results and inferences,[14] strategies for estimating the pattern of connectivity within empirical models in which *unobserved*

[11] In the case of time-variant $\mathbf{W}_{N,t}$ one must also fix $\mathbf{W}$ to some desired set (or path) of long run values.

[12] Another way to write (11) is $\mathbf{y}_t = [\mathbf{I}_N - \{\rho/(1-\phi)\}\mathbf{W}_{N,t}]^{-1}\{1/(1-\phi)\}(\mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t)$. This formulation highlights that, to get the steady state responses, one simply weights the parameters ρ and $\boldsymbol{\beta}$ in the immediate spatial effect calculation by the temporal multiplier (as emphasized, e.g., in [50]).

[13] In C & IPE specifically, and in research involving TSCS data sets (i.e., relatively small N, large T) more generally, one is frequently interested in unit-specific counterfactuals (as in our illustrative application, for example). In analyses of data sets with far larger numbers of units, especially when anonymously sampled, unit-specific responses will be of far less interest, and one is likely to prefer scalar summary average effect calculations (as suggested in [50], for instance) to vectors of unit-specific ones.

[14] Franzese and Hays [25,26] demonstrate analytically and in simulations, for instance, that relative misspecification of $\mathbf{W}$ tends to attenuate estimates of ρ while inflating those of the $\boldsymbol{\beta}$ on the $\mathbf{x}$'s most correlated with the true pattern of interdependence (i.e., with $\mathbf{Wy}$).

patterns of interconnections among units affect outcomes in those units have long interested spatial econometricians greatly. However, progress has been limited with regard to one or more central interests in social science research into interdependence: distinguishing the potential sources of spatial association—contagion, common exposure, or selection; testing or gauging the relative strengths of alternative mechanisms of contagion; calculating substantive spatiotemporal counterfactual effects; and allowing, or even potentially endogenizing, time-varying connectivity patterns.

Kooijman [47] and Openshaw [52] first suggested inductively choosing **W** to maximize Moran's I for **y** or to optimize univariate spatial lag models of **y**. Modern, sophisticated approaches following in that spirit include Boots and Dufournaud's [7] construction of binary contiguity matrices by marking dyads as contiguous or not to maximize the spatial association along that **W**; Griffith's [33–35] and Getis and Griffith's [31] spatial filtering strategies for extracting factors that maximally purge spatial association from **y**; Getis and Aldstadt's [29,30] similar aim, by a different route,[15] to separate the "variables being adjusted for spatial effects" into one component subject to spatial lag autocorrelation and one not; or Fernandez-Vazquez and Rodriguez-Valez's [16] alternative tack, noting that, assuming row normalization and exclusively non-negative w_{ij}, the rows of **W** can be seen as probability distributions for maximum entropy estimation of w_{ij} simultaneously with spatial lag parameter, ρ, and covariate coefficients, $\boldsymbol{\beta}$. Along another strategic line, Conley [12] proposed an instrumental variable (IV) approach to enhance robustness to errors in measured distances, d_{ij} (or in their mapping to w_{ij}). Chen and Conley [11] and Pinkse et al. [55] extend this line, adopting semiparametric approaches that approximate the unknown function mapping the distance metric(s), d_{ij}, to a univariate w_{ij}, by some flexible sequence of parametric families (e.g., polynomials in d), which they consistently estimate by IV or by sieves [32]. Bhattacharjee and Jensen-Butler [6] go further, building from Meen [51] to estimate $\rho\mathbf{W}$ without pre-specifying a distance metric, i.e., non-parametrically, directly from the spatial association of residuals from a first-stage non-spatial model. Our m-STAR approach resembles most that of Conley and Chen and Conley, excepting the IV projection of instruments onto a single dimension in estimating/instrumenting w_{ij}, which undermines any distinction of alternative mechanisms of interdependence. The non-parametric strategies in this vein share that limitation and furthermore must relegate any spatial dynamics exclusively to the error term because, as Anselin [2] first stressed, the first-stage non-spatial estimated errors are consistent only under spatial error and not spatial lag models.[16] The inductive strategies, as promising as they are toward their own aims, also exhibit one or more important limitations with regard to our list of research goals. In general, these limits can be summarized thus: univariate, pre-filtering, and/or non-parametric strategies tend to treat spatial effects or dynamics as nuisance to be purged or, at most, relegated to the error term, to treat all spatial association as arising from a single source, and/or to model association in ways that complicate or debar the extension to dynamic, endogenous interdependence and/or the consideration of the sorts of substantive counterfactuals that we seek. Our m-STAR strategy is simpler and parametric in its approach to these aims of distinguishing the potential sources of spatial association—contagion, common exposure, or selection; of testing or gauging the relative strengths of alternative mechanisms of contagion; of calculating substantive spatiotemporal counterfactual effects; and of allowing, or even potentially endogenizing, dynamic patterns of connectivity (networks).

In network analysis, on the other hand, estimation of the processes generating ties in *observed* networks, i.e., of selection, rather than the effects of alters' actions on egos' actions through the network, i.e., of contagion, is more usually the object of the study. Network models generally take the characteristics of units, including their behaviors, as *given, exogenous* explanators of which ties, typically seen in exclusively binary terms, will form between actors. From this vantage point, spatial

[15] The procedure finds a maximum distance, d_c, beyond which spatial correlation is negligible and sets weights w_{ij} in row i equal to a function of the [53] statistic of local association, $G_i^*(d_c)$, to 1, or to 0. Coefficients on dummies for rows with all $w_{ij} = 0$ then estimate the non-spatial component. Getis and Aldstadt [30] write a genetic algorithm for optimizing the procedure.

[16] Moreover, at least the current implementations of these approaches: identify the joint estimate of the interdependence pattern and strength from observed spatial association in part by assuming that $\rho\mathbf{W}$ fully captures all spatial association; they require symmetric and constant **W**; and they cannot separately identify ρ and **W** (our approach shares this last limitation).

econometric attempts to estimate and, ultimately, to endogenize w_{ij} within models of interdependent unit outcomes mirror network analytic attempts to model the *co-evolution* of unit behavior with the generation of network ties between units. In many social science research contexts, however, unlike the assumptions of extant network co-evolution models, relative connectivity is of degree rather than binary and, more dauntingly, the effective connectivity may not be directly observed.[17] Rather, commonly, one observes only some covariates theorized to relate to the effective connection. In C & IPE, for instance, such as in our application's context of interdependent ALM policymaking, 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.

In the network analytic tradition, Snijders and colleagues [61–63,65,66] have advanced perhaps furthest on this crucial next task in empirical modeling of dynamic, endogenous networks *cum* interdependence.[18] In this task, Snijders et al. [66] emphasize, as have we from a spatial econometric perspective, that the core challenge of disentangling the sources of network association (a.k.a., spatial correlation) is threefold. First, one must distinguish *influence* (i.e., *contagion* or *interdependence*) from *selection* (e.g., *homophily*) and from *social contexts* (i.e., common exposure) because, as we have shown elsewhere [20,21], any omissions or inadequacies in modeling those distinct sources of network or spatial correlation will bias inferences in favor of the included or better modeled mechanisms. Then, they also emphasize three further issues confronting such attempts: discrete time observation of continuous time processes, the need to control for alternative mechanisms and pathways by which observed networks and outcomes may arise, and the network dependence of the actors which precludes estimation by common statistical techniques, most of which assume independence.

To address these issues, they model the *co-evolution* of networks and behavior thus. N actors are connected by an observed, binary, endogenous, and time-variant matrix of ties, $\mathbf{x}$, with elements $\mathbf{x}_{ij}(t)$—in our notation, $\mathbf{W}_{NT}$, with elements $w_{ij,t}$. A vector of N observed, binary behaviors, $\mathbf{z}$, at time t has elements $\mathbf{z}_i(t)$—i.e., our $\mathbf{y}_t$ and y_{it}. Additional exogenous explanators may exist at unit or dyadic level, $\mathbf{v}_i(t)$ or $\mathbf{w}_{ij}(t)$—i.e., our $\mathbf{X}$. Opportunities arise for actors to change their network connections, switching at most 1 tie on or off, at continuous time fixed rate, λ_i^{net}, according to an exponential model. At present, λ_i^{net} is assumed constant for all i, j, t, though in principle one could parameterize it. Likewise, opportunities to increment, decrement, or leave the behavior arise at rate λ_i^{beh}.[19] When an opportunity to change network ties arrives for some i, i may choose to alter the status of any one of its $N-1$ ties to 'on' or to 'off' or to leave all ties unchanged. i makes these choices by comparing the values of some objective function of this form:

$$f_i^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z}) + \boldsymbol{\varepsilon}_i^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z}) = \sum_h \left\{\boldsymbol{\beta}_h^{net} \times \mathbf{s}_h^{net}(\mathbf{i}, \mathbf{x}, \mathbf{x}', \mathbf{z})\right\} + \boldsymbol{\varepsilon}_i^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z}), \tag{13}$$

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.[20] $s_h^{net}(\cdot)$ is some statistic, i.e., some function of the

[17] Both of these distinctions may result more from common simplifying assumptions in network analysis than from their underlying substance, however. Connections in friendship networks, for instance, are probably in truth more of relative degree than binary, and we often may not observe those friendship ties even as directly and imperfectly as by survey response indicating them or gauging their closeness.

[18] Leenders [48,49] presages. Hoff and colleagues [41–43] offer an alternative, Bayesian latent space approach. The Appendix offers a more detailed overview of the model of Snijders and colleagues. Our review follows [61,65,66] most specifically.

[19] Since observation occurs at discrete intervals, the freedom to vary these continuous time rates renders the assumptions of one actor making one unit-valued change in his/her network ties or behavior at a time 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.

[20] Snijders et al. ([66]:21) actually omit the stochastic component from the right hand side of (13), and seem to carry this omission into the simulation model implementation and the associated estimation. We suspect that this is highly consequential because it suppresses the dependence across units or dyads of their multinomial choices (see footnote 21) of which if any $\mathbf{x}_{ij}$ to switch on or off.

data, $\mathbf{x}, \mathbf{x}', \mathbf{z}$, that reflects the actor's objectives (ideally, substantively-theoretically derived) regarding to the network, $\mathbf{x}$, and behaviors, $\mathbf{z}$. The $\boldsymbol{\beta}_h^{net}$ to be estimated are the relative weights of these objectives. Assuming the $\boldsymbol{\varepsilon}_i^{net}(\mathbf{x}, \mathbf{x}', \mathbf{z})$ extreme value distributed, independently across actors (see footnote 20) and over time, yields the multinomial logit model of categorical choice. Similarly, when an opportunity to change behavior arrives, actor i compares the value of an analogous objective function under each of three possible actions: increment or decrement by 1 or leave unchanged. Formally, i compares $\mathbf{z}$ to $\mathbf{z}'$ given $\mathbf{x}$ and $\mathbf{z}_{j\neq i}$. Again assuming *i.i.d.* extreme value stochastic components, the multinomial logit emerges.

Identification of the model derives from a battery of strong conditional independence assumptions, namely: conditional on the existing vector of behaviors and set of network ties, each unit's choices on its network ties and on its behavior and its selection to act on its ties or behavior are all independent of each other and of all other units' choices and selections. Operationally, this amounts to conditioning on the first observation and assuming that simultaneity is not *effectively instantaneous* (as elaborated above) and that temporal precedence implies causal precedence.[21] Given this, estimation proceeds by simulating sequences of behaviors, $\mathbf{z}$, and networks $\mathbf{x}$, and searching over permissible values of the parameters, $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$, to minimize discrepancy from the observed sequences of $\mathbf{x}$ and $\mathbf{z}$ to the simulated sequences. Standard errors could then derive from jackknife or bootstrapped resampling [64] if explicit likelihoods or moment equations are unavailable for applying the standard analytic formulae.

Some features of this network co-evolution model, for all the remarkable advances it represents, are limiting (at least) for C & IPE. First, least critically, relative connectivity of the units and many behaviors or attributes of interest are less often binary or ordinal in C & IPE. In the canonical globalization and tax competition context, for instance, the outcomes of interest are tax rates, and many sources of interdependence will surround the (continuous) strength of economic competition. Second, more importantly, strengths of relative connectivity are often unobserved, or even unobservable, at least in C & IPE. Continuing the example, we observe only some covariates theorized to relate to the unobserved strengths of competition, such as geographic contiguity and proximity or trade and capital flows. Thus, the left hand side of the selection components of network co-evolution models would have no data. Indeed, one could estimate networks and their determinants *only* by estimating their impact on actors' behaviors, *given some model of how networks matter for that behavior and how the observed covariates relate to the unobserved network ties*. Third, temporal precedence often will not suffice to assure causal precedence in C & IPE contexts. Most political economy relations are strategic and, in strategic interactions, the effect of *alters* on *ego* can even be based on expected futures. The strategic interdependence of policies across units, for example, arises from policymakers' *simultaneous* play of a game in which each actor's optimal policy depends on others' current or expected future policies. Moreover, as previously noted, empirically practically, *simultaneous* means *within an observational period, etc.*, and many C & IPE contexts have high frequency behavior and/or network changes relative to far lower periodicity of observation. Consider capital flows versus annual budgets, for example. Finally, conditioning on the first observation is least consequential with large T, yet typical C & IPE data sets have only moderate T (e.g., 20–40 years)

21 Some identification issues seem to us to persist, notwithstanding these strong assumptions. For one, assuming independent multinomial decisions for the endogenous behaviors, network ties, and selections to act would seem to undo some of the allowance for dependence in those choices, although it yields the great advantage of perhaps validating standard multinomial logit and exponential hazard rate models for those system components. Another issue is that included among the conditioning 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 actors' multinomial choices. Heckman's [40] critique of including the actual outcomes on the right hand side of such latent variable models may apply. If so, including latent variables or estimated probabilities instead as then required would imply stochastic component dependence across units, rendering these multinomial logits N dimensional rather than simple unidimensional optimization exercises. However, beyond some conjectures we offer in the conclusion, we have little progress on those issues to offer here, nor do we know of any other scholarship that has made greater progress on these issues in this behavior–network co-evolution context. Spatial econometricians have made considerable progress on this multidimensional optimization issue of interdependent latent variable models (and we have also discussed some of these options for political scientists: Franzese and Hays [22,24–26], Hays [36], Hays and Kachi [39]), but entirely outside of the co-evolution context.

and longitudinal network data sets only very small T (less than 5 periods, even just two). Obviously, to a considerable extent, these are all general concerns, not necessarily restricted to C & IPE.

As Leenders ([49]:165) also notes: most research on network/spatial dependence either studies the formation of networks (*selection*), taking unit attributes and behaviors as given, or the effects on units' behaviors of network interdependence (*contagion*), taking patterns of connectivity as given. Spatial econometricians have primarily worked in the latter mode; network analysts have primarily worked in the former; but both are eager to combine the two. Although other differences in tendency appear to us – e.g., 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 – the difference in core question – what explains networks vs. how does network dependence affect outcomes – seems to us the most crucial. Among network analysts, co-evolution models like that just reviewed represent the greatest advances toward combining of *contagion* and *selection*.

Our approach comes instead from a spatial econometric vantage point, so it begins with the spatiotemporal lag model, (1), and seeks to expand its specification to allow estimation of $\mathbf{W}$, the matrix of relative connectivity, a.k.a. the network, i.e., to model 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 *selection* from the view and in the language of network analysts. For instance, the sociologists' *homophily* ("like seeking like"), if based on exogenous characteristics of ego and alter, parallels, from the spatial econometric view, a model of w_{ij} as a function of some $\mathbf{x}_i$ and $\mathbf{x}_j$. If we consider (some function of) the vector of behaviors of interest, $\mathbf{y}$, among these explanators of $\mathbf{W}$, this parallels a stronger form of homophily selection in which network ties and actors' behaviors are jointly endogenous, raising greater statistical challenges. Thus, the STAR model with estimated, endogenous spatial weights is the spatial econometric analogue to the network co-evolution model, integrating contagion and selection. Now consider m-STAR, i.e., a multiple-spatial-weights STAR model, from this vantage point:

$$\begin{aligned}\mathbf{y} &= \rho_1\mathbf{W}_1\mathbf{y} + \rho_2\mathbf{W}_2\mathbf{y} + \cdots + \rho_R\mathbf{W}_R\mathbf{y} + \phi\mathbf{M}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ &= \mathbf{W}\mathbf{y} + \phi\mathbf{M}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \text{where } \mathbf{W} \equiv \sum_{r=1}^{R} \rho_r\mathbf{W}_r, \end{aligned} \tag{14}$$

or, in scalar notation,

$$\begin{aligned} y_{i,t} &= \rho_1 \sum_j w^1_{ij,t} y_{j,t} + \rho_2 \sum_j w^2_{ij,t} y_{j,t} + \cdots + \rho_R \sum_j w^R_{ij,t} y_{j,t} + \phi y_{i,t-1} + \sum_k x^k_{i,t}\boldsymbol{\beta}_k + \boldsymbol{\varepsilon}_{i,t} \\ &= \sum_j \left(\rho_1 w^1_{ij,t} + \rho_2 w^2_{ij,t} + \cdots + \rho_R w^R_{ij,t}\right) y_{j,t} + \phi y_{i,t-1} + \sum_k x^k_{i,t}\boldsymbol{\beta}_k + \boldsymbol{\varepsilon}_{i,t} \\ &= \sum_j \left\{ \left(\sum_r \rho_r w^r_{ij,t} \right) y_{j,t} \right\} + \phi y_{i,t-1} + \sum_k x^k_{i,t}\boldsymbol{\beta}_k + \boldsymbol{\varepsilon}_{i,t}. \end{aligned} \tag{15}$$

As line 2 of (15) perhaps best clarifies, the term in parentheses is a parameterized (linear additive) model of the weights on $y_{j\neq i}$ in affecting y_i, with the w_{ij} being the covariates expected to explain the pattern and relative strength of interdependence and the ρ_r being their coefficients to estimate. Thus, we can conceive the m-STAR model as a spatiotemporal lag model with estimated $\mathbf{W}$, with $\hat{\mathbf{W}} = \sum_r \hat{\rho}_r\mathbf{W}_r$ being a weighted sum of the $\mathbf{W}_r$, which are the covariates thought to explain net connectivity. (Note the scaling issue though: the interdependence strength and the pattern in which it occurs are not separately, but rather only jointly, identified in this model of *implicit net network interdependence*.) If, furthermore, any $\mathbf{W}_r$ has elements that are functions of $\mathbf{y}$, then $\mathbf{W}$ and $\mathbf{y}$ are jointly endogenous, making (15) a co-evolution model.

The sorts of models of $\mathbf{W}$, i.e., of networks, expressible thus are limited, without considerable further complication, to those with continuous relative strengths of ties, w_{ij}. For truly binary connectivity, one would need to transform the term in parentheses by, say, applying the log odds function and a decision rule to convert probabilities to (0, 1). (If one believes, as we tend to do, that

connectivity is a degree, measured at best with error, this is no limitation.) These or other non-linear additive models of w_{ij} would also generally entail multidimensional integration complications (see footnote 21). Then, too, the cost in estimation complexity of enriching the model of connectivity by adding $\mathbf{W}_r$ covariates is high compared to that of adding unit, dyad, or exogenous external covariates to the $\mathbf{X}\boldsymbol{\beta}$ term (but perhaps not compared to similarly enriching network co-evolution models). The approach has some major advantages too though, notably among them simplicity and that fully developed likelihoods for the m-STAR model exist, at least in the exogenous $\mathbf{W}_r$ case, both the simpler one that conditions on the first observation and the unconditional one better suited for *instantaneous* interdependence or small T. Thus, we can apply all the apparatus for estimating spatial econometric models and all the procedures for interpreting their spatiotemporally dynamic effects discussed above. Conversely, we can use all the standard network analytic tools interpret and present the estimated *implicit net network interdependence*, $\hat{\mathbf{W}}$, as we shall see.

The conditional likelihood for m-STAR extends (4) for STAR:

$$\ln L(\rho, \phi, \boldsymbol{\beta}, \sigma; \mathbf{y}, \mathbf{X}) = \ln(2\pi\sigma^2)^{-NT/2} + \ln|\mathbf{A}| - \frac{1}{2\sigma^2}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}, \tag{16}$$

where

$$\mathbf{A} = \mathbf{I}_{NT} - \mathbf{W} \quad \text{and} \quad \boldsymbol{\varepsilon} = \mathbf{A}\mathbf{y} - \phi\mathbf{M}\mathbf{y} - \mathbf{X}\boldsymbol{\beta}.$$

The unconditional likelihood extends (6) analogously. The estimated variance of $\hat{\mathbf{W}}_{ij}$ is

$$\hat{\mathbf{W}} = \sum_r \hat{\rho}_r \mathbf{W}_r \Rightarrow \widehat{\mathrm{Var}(\hat{\mathbf{W}}^{(i,j)})} = \left[\mathbf{W}_1^{(i,j)}\mathbf{W}_2^{(i,j)} \cdots \mathbf{W}_R^{(i,j)}\right] \hat{\boldsymbol{\Omega}}_{\hat{\boldsymbol{\rho}}} \left[\mathbf{W}_1^{(i,j)}\ \mathbf{W}_2^{(i,j)} \cdots \mathbf{W}_R^{(i,j)}\right]', \tag{17}$$

where $\hat{\boldsymbol{\Omega}}_{\hat{\boldsymbol{\rho}}}$, the estimated variance–covariance matrix of $\hat{\boldsymbol{\rho}}$, is minus the inverse of the Hessian of the likelihood in the usual fashion. Written in $(N \times 1)$ vector notation, the conditional likelihood for the m-STAR is again mostly conveniently separable into parts:

$$\begin{aligned}\mathrm{Log} f_{\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{y}_2 | \mathbf{y}_1} &= -\frac{1}{2}N\,(T-1)\log\left(2\pi\sigma^2\right) + (T-1)\sum_{r=1}^{R}\log\left|\mathbf{I}_N - \rho_r\mathbf{W}_{N,r,t}\right| \\ &\quad - \frac{1}{2\sigma^2}\sum_{t=2}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t \quad \text{where } \boldsymbol{\varepsilon}_t = \mathbf{y}_t - \sum_{r=1}^{R}\rho_r\mathbf{W}_{N,r,t}\mathbf{y}_t - \phi\mathbf{y}_{t-1} - \mathbf{X}_t\boldsymbol{\beta},\end{aligned} \tag{18}$$

and the unconditional (exact) likelihood analogously extends (6). In our example, T is large enough that the more compact conditional likelihood is acceptable.

Co-evolution models, i.e., models where network connectivity, $\mathbf{W}$, is some function of $\mathbf{y}$, present larger challenges. Our simple stratagem for a first cut is to invoke the same *poor man's exogeneity* as is assumed in network co-evolution models (and widely elsewhere): i.e., temporally lag the $\mathbf{y}$ in this function explaining $\mathbf{W}$ and assume that the conditions required for this identification approach hold sufficiently well. As just argued, however, this does not address the problem of true or effective simultaneity, which seems a likely situation for C & IPE contexts at least. We will therefore propose in the conclusion a two-step estimation procedure addressing true simultaneity for future exploration.

4. Empirical illustration

The Appendix demonstrates analytically and through simulations that the S-ML estimator performs well in absolute terms and greatly outperforms simpler least squares estimators for these m-STAR models. Our findings there parallel those from our earlier studies [17,18,21,22,24]—notably, in highlighting how:

1. Omission (or relative misspecification) of common exposure, contagion, or selection sources of network/spatial association induces overestimation of included (or better specified) sources, with the biases distributed across the included in proportion to how similar are the spatial association patterns induced by the omitted and included sources;

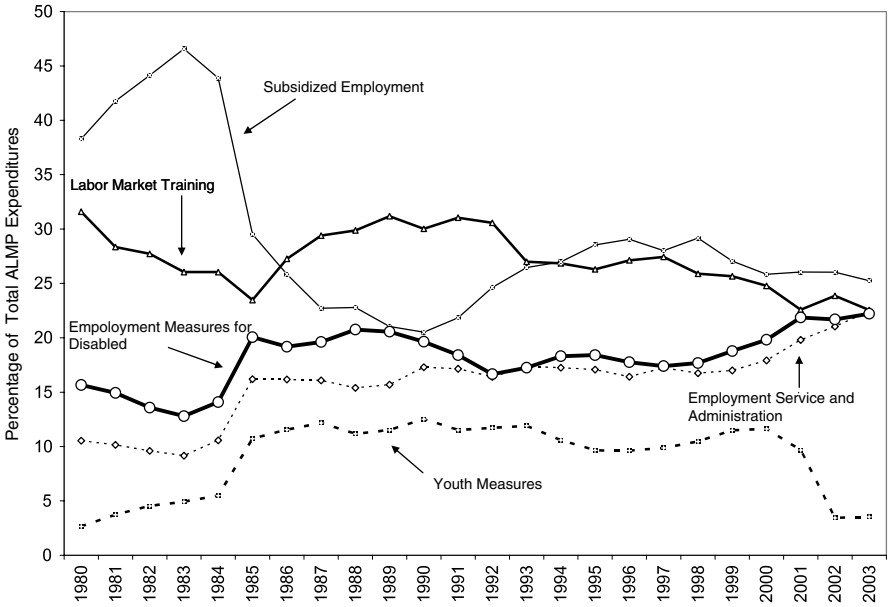

Fig. 1. Disaggregated active labor market expenditures in the OECD, 1980–2003.

2. The simultaneity bias of least squares generally inflates estimates of contagion (and/or of selection, whichever is stronger), inducing deflation bias in other sources, again in proportion to the similarity of spatial association patterns induced by these sources.

So we proceed here in the text directly to our illustrative application.

To illustrate application of our m-STAR approach (with first-cut identification strategy of time lagging assumed adequate), we extend our previous analysis of active labor market (ALM) policies [19]. First, we extend the sample to include both *EU* and non-*EU* states[22] to allow distinction of interdependence among *EU* members due to co-membership from global interdependence. Intra-EU diffusion would suggest positive contagion and intra-EU free-riding, negative. Next, as argued in [19], geographic contiguity may induce free-riding (and negative contagion) among bordering states, primarily *via* employer location decisions.[23] Lastly, economic competition suggests competitive races (to bottom, mean, top, or elsewhere), implying positive ALM policy interdependence in proportion to the economic interchange between states.

OECD data sets give ALM program expenditures by five categories: labor market training (LMT), youth measures, subsidized employment (SEMP), public employment services and administration, and disability measures. Fig. 1 plots *OECD* average spending by type over time. SEMP and LMT are the two largest components over the entire sample period, accounting for 26.9% and 26.7% total expenditures.

Table 1 gives the programmatic breakdown in ALM expenditures by state, revealing great variation across the sample. Sweden and Denmark were big spenders *per capita* ($360.88 and $287.20 (2000, PPP$)); the US and Greece spent least ($43.72; $34.97). Table 1 also reveals some spatial clustering: the

22 Annual 1980–2003 data for 21 OECD countries (of which 8 are EU members at the start, rising to 14 *by the end*): Australia, *Austria, Belgium*, Canada, *Denmark, Germany, Greece, Finland, France, Ireland, Italy*, Japan, New Zealand, the *Netherlands*, Norway, *Portugal, Spain, Sweden*, Switzerland, the UK, and the US.

23 The key mechanism seems to be this: effective ALM policies enhance attractiveness to firms of locating production in the spending economy, which, because firms tend to locate at borders to minimize transport and transaction costs while maximizing "comparative advantages" (or tax evasions), in turn enhances the appeal of locating in bordering states. Corroborating evidence is in [54], inter alia.

Table 1
Disaggregated active labor market expenditures per capita (2000 PPP$).

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

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

Scandinavian states spent much more than the OECD average; the Portuguese and Spanish averages over this period are within $1; and Australia, New Zealand, Canada, and the US spent well below average. What explains these patterns: strategic policy interdependence, similar exogenous external conditions, correlation of domestic factors, or some as yet unconsidered selection process?

We estimate an m-STAR model with co-evolutionary dynamics to address this question.[24] The model, in matrix notation, is

$$\mathbf{y} = \left[\sum_{r=1}^{R} \rho_r \mathbf{W}_r\right] \mathbf{y} + \phi \mathbf{M}\mathbf{y} + \gamma \mathbf{L}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{19}$$

where $\mathbf{y}$, ALM expenditures, is an $NT \times 1$ vector of cross-sections stacked by periods as described before. The rth spatial autoregressive coefficient is ρ_r, and $\mathbf{W}_r$ is an $NT \times NT$ block diagonal spatial weighting matrix. Each $\mathbf{W}_{r,t}$ contains elements $w^r_{ij,t}$ reflecting a particular basis from which we expect interdependence in ALM policy to arise: geographic proximity, *EU* co-membership, and economic interdependence. These are the exogenous explanators of *net network interdependence*: $\hat{\mathbf{W}} = \sum_r \hat{\rho}_r \mathbf{W}_r$. In the other new term, γ is the co-evolutionary dependence parameter, and $\mathbf{L}$ is an $NT \times NT$ "policy distance" matrix with $|y_{i,t-1} - y_{j,t-1}|$ in cells (it, jt). The addition of $\gamma\,\mathbf{Ly}$, therefore, reflects a substantive proposition that states with more similar ALM policies (specifically, spending levels) affect each other's ALM policies more ($\gamma < 0$), as in the network analyst's *homophily*, or less ($\gamma > 0$, *heterophily*), than do states with less similar policies. It is the endogenous determinant of the *implicit net network* in our conceptualization. The rest is as before: $\mathbf{My}$ is the first-order time lag, ϕ its coefficient; $\mathbf{X}\boldsymbol{\beta}$ are the exogenous non-spatial components and $\boldsymbol{\varepsilon}$ the disturbances assumed *i.i.d.*

$\mathbf{L}$ contains lagged y's, so (19) is a system of N equations non-linear in the endogenous variable. This greatly complicates spatiotemporal dynamics. Without linear multipliers or analytic solutions

24 Case et al. [10], Brueckner and Saavedra [8], Fredriksson and Millimet [27], Redoano [57], Allers and Elhorst [1] among others have used spatial lag models to test strategic policy interdependence hypotheses, but none use multiple spatial lags or consider co-evolution.

for steady states, we must calculate spatiotemporal co-evolutionary responses to changes in $\mathbf{X}$ or $\boldsymbol{\delta}$ recursively. To start, rewrite (19) as t cross-sections:

$$\mathbf{y}_t = \mathbf{W}_{N,t}\mathbf{y}_t + \phi\mathbf{y}_{t-1} + \gamma\left[abs\left(\boldsymbol{\Pi}\left[\mathbf{y}_{t-1} \otimes \mathbf{I}_N\right]\right)\right]\mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \tag{20}$$

where $\mathbf{y}_t$, $\mathbf{W}_{N,t}$, and $\mathbf{X}_t$ are $N \times 1$, $N \times N$, and $N \times 1$ matrices, and $\boldsymbol{\Pi}$ is an $N \times N^2$ matrix produced by horizontally concatenating N separate $N \times N$ matrices. The ith $N \times N$ matrix in $\boldsymbol{\Pi}$ has -1's on its diagonal and 1's for each element of the ith column, except for element (i, i) which is 0, as are all remaining elements. For instance, with $N = 3$,

$$\boldsymbol{\Pi} = \begin{bmatrix} 0 & 0 & 0 & -1 & 1 & 0 & -1 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & -1 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{21}$$

The function *abs* is element by element absolute value; its argument gives the vector of differences $y_{i,t-1} - y_{j,t-1}$, reflecting policy distance. In reduced form, (20) is

$$\mathbf{y}_t = \left[\mathbf{I}_N - \mathbf{W}_{N,t} - \gamma\left[abs\left(\boldsymbol{\Pi}\left[\mathbf{y}_{t-1} \otimes \mathbf{I}_N\right]\right)\right]\right]^{-1}\left(\phi\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right). \tag{22}$$

Our empirical analysis estimates (19) and applies (22) to study total ALM spending and its two largest components, subsidized employment (SEMP) and labor market training (LMT). Our dependent variables are measured per capita (2000, PPP$), and the key right hand side variables, which allow us to evaluate the spatial interdependence among our sample states, are the various spatial lags of ALM spending. In detail, our spatial lags, $\mathbf{W}_r\mathbf{y}$, use four different weights matrices ($R = 4$). $\mathbf{W}_1$ is a *binary contiguity weights matrix*, coding $w_{ij} = 1$ for states i and j that share a border and $w_{ij} = 0$ for states that do not, with exceptions France, Belgium, and Netherlands treated as contiguous with UK, Denmark with Sweden, and New Zealand with Australia. $\mathbf{W}_2$ is an EU co-membership weights matrix; i.e., $w_{ij} = 1$ if both i and j are EU members and $w_{ij} = 0$ otherwise. $\mathbf{W}_3$ has weights that reflect the trade (imports + exports) shares between sample states,[25] and $\mathbf{W}_4$ is the policy distance co-evolution matrix described above, with $|y_{i,t-1} - y_{j,t-1}|$ in cells (it, jt). For estimation, we *row standardize* (as commonly done in spatial econometrics research) all four matrices by dividing each cell in a row by that row's sum.[26]

We also include several domestic variable controls, starting with macroeconomic performance as gauged by the real GDP growth and unemployment rates. Governments in wealthier economies tend to provide more public goods and services (Wagner's Law), suggesting a positive coefficient estimate for real GDP. Alternatively, Baumol's Disease, which refers to how decreasing relative productivity in service sectors renders financing of public services increasingly costly as economies grow and shift toward service sector activity, suggests a negative relation of wealth to spending. GDP *growth*, though, should mostly capture pseudo-automatic responses to macroeconomic conditions, and so get a negative coefficient. We expect positive unemployment coefficient for the same reason.

Next, we control three structural features of a state's economy related to its labor markets and exposure to external shocks. The labor market factors are union density and Iversen and Cusak's [44] deindustrialization measure. Higher union density boosts the influence of organized labor, so we expect it to have positive coefficient. Iversen and Cusak [44] argue that workers cross significant

25 More exactly, $w_{ij,t}$ sums exports i to j and j to i and imports i from j and j from i and divides by 4 times i's GDP. We use all four values to average across the slight discrepancies in the data. Our trade flow $\mathbf{W}$ may raise concerns for some about the endogeneity of those $w_{ij,t}$ or the possibility of *strong connectedness* among our units. We do not view either as a serious practical issue here. First, given row standardization, our weights reflect relative shares of trade across partners. Insofar as ALM programs may stimulate the economy, they would increase trade with all partners, but we do not expect states' ALM spending to affect dyadic trade *shares* noticeably. (For example, US ALM spending likely explains extremely little of why the US trades more with Canada than Austria.) Second, note that *strongly connected* $\mathbf{W}$ are those in which all the off-diagonal elements are close to their maximum values. Our trade flow $\mathbf{W}$ has all non-zero, but not remotely all near maximum, elements. Moreover, insofar as a *strong connectivity* bias may arise, with positive interdependence (as found here), the bias is downward [60]. That is, *strong connectivity* would likely induce underestimation of the interdependence generated by trade ties.

26 Row normalization is not necessarily substantively neutral (see, e.g., [56]).

skill barriers in shifting to services from manufacturing and agriculture. Thus, we expect deindustrialization to spur LMT also. Many argue that global economic exposure spurs government spending, especially on programs that help workers adjust to external shocks (e.g., [59,9,45,58,38,37]).[27] We measure exposure by trade openness.

We also include the working age share of the population, the share of cabinet seats held by left and Christian Democratic parties, and the share 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 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., [28,67,46]). The simpler left–right ideological dimension may also relate to ALM programs.

Table 2 presents our results. We estimate three regressions for each dependent variable: the first is non-spatial with country indicators, the second spatial with country indicators, and the third spatial with country and time indicators. Using period dummies is a flexible way to model common OECD-wide trends and/or common (exogenous) shocks in ALM expenditures. Recall that the most important issue methodologically in obtaining good estimates of the strength of interdependence, i.e., of $\boldsymbol{\rho}$, is to model well any alternative mechanisms by which the outcomes might correlate spatially, such as common exogenous shocks (e.g., global economic conditions) or correlated domestic factors (this is *Galton's Problem*). From that perspective, the set of country and year dummies serve as a powerfully conservative way to account for common (or correlated) outside shocks or spatially correlated (near common) domestic factor. Failure to account for such alternatives will bias spatial lag coefficient estimates, usually positively.[28]

The non-spatial and spatial model estimates support differing explanations for total ALM expenditures. The non-spatial model points to domestic real GDP growth, oddly indicating a strong *pro*cyclicality to total ALM (although not to its LMT or SEMP components), and to the labor market structures of deindustrialization and especially union density. The spatial models suggest that the effects that the non-spatial models attribute to domestic growth and trade exposure instead reflect spatial diffusion of responses to global conditions. The odd procyclicality in ALM policy fades or reverses to the more plausible countercyclicality in the spatial models. Also interestingly, all three models find sizable differences in sources of LMT versus SEMP spending. LMT seems closely related to workforce age demographics and not very closely related to our labor market structural or institutional measures. SEMP, in contrast, 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, notably countercyclical to the domestic economy. Perhaps most interestingly, the spatial models suggest that, while total ALM spending is not especially partisan, its composition is decidedly so, with LMT associated positively and SEMP negatively with left cabinets. More centrally for us, finally, Wald tests of the spatial lag coefficients reveal strong evidence of interdependence in ALM policy, the t-tests on 13 of the 24ρ_r estimates being significant at conventional levels and the six joint tests of the 4ρ_r estimates *per* spatial model all overwhelmingly rejecting null hypotheses of zero coefficients, i.e., of the non-spatial model.[29] In particular, total ALM spending seems strongly spatially interdependent along all four dimensions of proximity, SEMP much less so, and LMT intermediately. Consequently, coefficient estimates in non-spatial models, especially of total ALM, will almost certainly suffer from omitted-variable bias.

We focus, therefore, on the spatial models of total ALM spending and, in particular, on the most conservative time dummies model 3. In this model, we see few strong and significant

27 Others stress how increased international exposure produces competitive pressures favoring smaller governments, but this is more properly reflected by spatial lags, as in our $\mathbf{W}_3\mathbf{y}$ term here (see, e.g., Basinger and Hallerberg [4] or Franzese and Hays [21,23]).

28 See the Appendix for background on *Galton's Problem* and for analytic and simulation results that support the continued centrality of its redress to valid and accurate inference in the m-STAR context.

29 Likelihood ratio tests of the models and information criteria also strongly favor the spatial models.

Table 2
ALM spending models — estimation results.

DEP. VAR. →	Total ALM			LMT			SEMP		
INDEP. VAR. ↓	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	−195.89*	−150.369	51.487	−162.53***	−160.30**	−100.950	−122.27**	−129.70**	−65.945
	(103.645)	(114.470)	(125.395)	(59.486)	(64.481)	(69.774)	(57.794)	(61.645)	(71.708)
Temporal lag	0.875***	0.836***	0.801***	0.872***	0.865***	0.758***	0.830***	0.848***	0.704***
	(0.026)	(0.032)	(0.0407)	(0.028)	(0.031)	(0.040)	(0.034)	(0.044)	(0.062)
Real GDP growth rate	1.365***	0.449	−1.100	0.269	0.237	−0.861	0.036	−0.056	−1.044*
	(0.411)	(0.607)	(1.144)	(0.201)	(0.258)	(0.630)	(0.188)	(0.231)	(0.629)
Stdzd unemployment rate	−0.070	0.552	0.125	0.361	0.844*	0.541	−0.169	−0.351	−0.544
	(0.791)	(0.835)	(0.876)	(0.445)	(0.456)	(0.509)	(0.438)	(0.448)	(0.490)
Union density	0.888***	0.711**	0.527*	0.205	0.138	0.259	0.567***	0.645***	0.629***
	(0.302)	(0.318)	(0.321)	(0.167)	(0.176)	(0.176)	(0.169)	(0.175)	(0.184)
Deindustrialization	1.259	1.209	0.499	0.106	0.105	−0.183	1.426***	1.440***	1.0514**
	(0.785)	(0.776)	(0.773)	(0.449)	(0.447)	(0.435)	(0.438)	(0.438)	(0.440)
Trade openness	−0.522***	−0.183	−0.192	−0.187*	−0.083	−0.009	0.018	0.039	0.104
	(0.169)	(0.196)	(0.213)	(0.097)	(0.114)	(0.118)	(0.104)	(0.113)	(0.133)
Working age population	0.946	0.216	0.108	2.239***	2.113**	2.623***	−0.139	−0.075	0.189
	(1.498)	(1.603)	(1.629)	(0.863)	(0.909)	(0.940)	(0.829)	(0.860)	(0.934)
Left cabinet seats	−0.024	0.001	0.001	0.035	0.044**	0.057***	−0.049**	−0.047**	−0.037*
	(0.039)	(0.039)	(0.038)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Christian Dem. cab. seats	−0.160*	−0.102	−0.085	−0.070	−0.074	−0.042	−0.032	−0.031	−0.037
	(0.095)	(0.096)	(0.093)	(0.054)	(0.053)	(0.052)	(0.053)	(0.053)	(0.052)
Left libertarian vote	−0.285	−0.549	−0.728	−0.316	−0.535	−0.714**	−0.248	−0.208	−0.147
	(0.622)	(0.611)	(0.598)	(0.355)	(0.361)	(0.352)	(0.346)	(0.344)	(0.340)

Table 2 (*continued*)

DEP. VAR. →	Total ALM			LMT			SEMP		
INDEP. VAR. ↓	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Spatial weights:									
Borders		−0.112***	−0.098***		−0.185***	−0.188***		0.058	0.054
		(0.040)	(0.040)		(0.057)	(0.060)		(0.054)	(0.056)
EU membership		−0.071***	−0.060*		−0.079	−0.183**		−0.058	−0.097
		(0.034)	(0.036)		(0.061)	(0.073)		(0.062)	(0.065)
Trade shares		0.386***	0.239**		0.316***	0.180		−0.017	−0.097
		(0.101)	(0.111)		(0.119)	(0.159)		(0.148)	(0.169)
Policy distance		−0.053	−0.189***		−0.004	−0.298***		0.049	−0.205**
		(0.039)	(0.067)		(0.046)	(0.085)		(0.049)	(0.087)
TIME DUMMIES?	*No*	*No*	*Yes*	*No*	*No*	*Yes*	*No*	*No*	*Yes*
σ	21.221***	20.741***	19.564***	12.061***	11.828***	11.161***	11.794***	11.702***	11.154***
	(0.787)	(0.775)	(0.722)	(0.454)	(0.451)	(0.426)	(0.437)	(0.434)	(0.414)
Log-likelihood	−1646.41	−1638.99	−1616.95	−1438.49	−1432.5	−1413.61	−1430.23	−1427.63	−1410.92

All regressions include fixed state 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 .10 level.

** Significant at the .05 level.

*** Significant at the .01 level.

effects of domestic conditions net of interdependence, common shocks, and fixed state-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 response to domestic growth are close to significant and sizable. No response 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 the composition of ALM spending. Regarding the estimated pattern of interdependence, i.e., the *implicit net network*, we find interdependence strongly negative among bordering states and moderately negative among EU states. The sign and relative strengths of interdependence obtained from these patterns are consistent with our [19] argument regarding positive externalities inducing free-riding. The negative EU co-membership $\hat{\rho}$ also bolsters the case for those concerned that the EU is not adequately fostering coordination in employment policy. The positive coefficient for the trade weights spatial lag, conversely, supports arguments of globalization-induced competition. The coefficient on the (endogenous) policy distance spatial lag is negative and quite significant (for total ALM, and for LMT and SEMP as well). This indicates lesser dependence of domestic ALM spending on states whose spending differs more from the domestic level; policymakers follow more closely those more similar to them, as revealed by the similarity in their policies: *homophily* in network analytic terms.

We are satisfied that ALM policy exhibits statistically significant interdependence, and that patterns of interdependence relate to geographic contiguity and to EU co-membership in ways that indicate policy free-riding, to trade relations in a manner implicating globalization and policy competition, and to policy distance so as to suggest homophily, but what do these results tell us of the net sign and substantive magnitude of this implicit net network or of the contagion effects of some states' ALM policies on policymakers in other states *via* this estimated network? What do they say about the ALM policy responses across these interdependent political economies to hypothetical variations in domestic and/or foreign conditions or policies? Answers and fuller interpretation of the co-evolutionary spatiotemporal effects and dynamics that these estimates imply require the spatial multiplier given in (22).

The non-linear spatial multiplier in (22) captures feedback from, say, Belgium on France and other states, and back from France and those others onto Belgium, and so on. The immediate time t effect on policies in all 21 states, $\mathbf{y}_t$, from a given set of time $t-1$ policies, $\mathbf{y}_{t-1}$, including spatial feedback, can be calculated with this spatial multiplier for any desired counterfactual shocks to the rest of the right hand side of (22). To find the LRSS (cumulative) level of $\mathbf{y}$, we must solve (22) recursively. With exogenously time-varying $\mathbf{W}$, like our trade weights, $\mathbf{W}_3$, we also need to specify the values of $\mathbf{W}_3$ that we assume to obtain in the period considered or in the long run. With endogenously time-varying $\mathbf{W}$, like our policy distance matrix $\mathbf{W}_4$, the system is much more complex. (Stationarity, e.g., must not only obtain initially but also continue to hold as dynamics unfold.) To get variance estimates, once again, we could use the delta method or simulate them by a parametric bootstrap of the procedure described above. Given the considerable non-linearity of (22), the simulated standard error estimates may have better properties; they are certainly far easier to generate.

Table 3 illustrates calculation of the estimated co-evolutionary spatiotemporal responses to hypothetical shocks. For this example, we start with the 2001 values for all exogenous variables (the last year for which all the states have data) and, using the parameter estimates from model 3 of Table 2, determine the LRSS levels of expenditures by recursive calculation of (20). Then, with the system at this steady state, we shock each state's ALM spending by \$1 (i.e., a \$1 shock to $\boldsymbol{\delta}$) and calculate the new LRSS. These results show the estimated system to be reasonably stable and these LRSS effects sizable. The specific pattern found suggests convergence, with previously low/high spenders spending more/less in the new steady state.

We can also use (20) to plot estimated co-evolutionary spatiotemporal paths of responses to hypothetical shocks. Again using the 2001 values of the exogenous variables and starting from the LRSS that would emerge from those values and the parameter estimates, Fig. 2 plots as an example the ten-period response in German, Austrian, and French ALM spending to a \$10 permanent positive shock to Germany's ALM spending ($\boldsymbol{\delta}$). Note the differences in how France and Austria respond to the German shock, Austria converging toward Germany's permanently higher ALM spending and

Table 3
Effects of a common counterfactual ($1) shock to ALM spending.

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

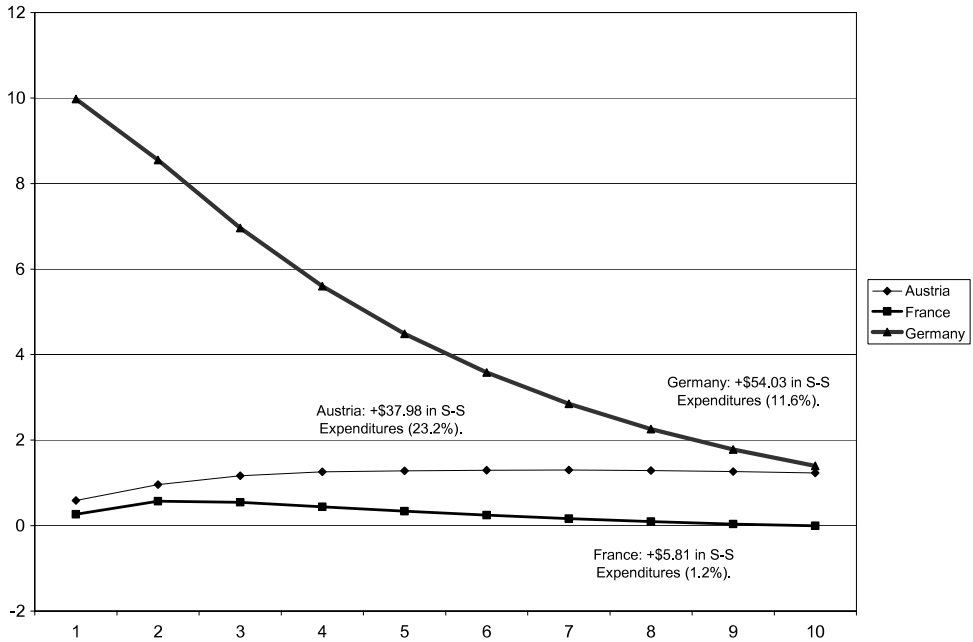

Fig. 2. Co-evolutionary spatiotemporally dynamic response of German, Austrian, and French ALM spending to a +$10 permanent shock to German ALM spending.

France returning to its *status quo ante*. These reflect the latent structure of interdependence between these three states, which can be illustrated as exemplified in the network table and graphic shown next.

Table 4
Estimated ALM policy interdependencies, i.e., the net ALM policy network, in 1981.

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

Dependent variable: total ALM spending; model: 3. Actual weights multiplied by 10 (and standard errors adjusted accordingly) to improve table formatting.
* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

Using (17), we can also show the estimated weights matrix, i.e., the estimated pattern of *implicit net network interdependencies* in ALM policy among these states, as Table 4 does for the estimated network using the 1981 covariate values.[30]

We can also depict these estimated patterns and strengths of interdependence using graphical methods familiar in network analysis as, e.g., Fig. 3 does for the estimated *implicit net ALM policy network* in 1991.[31] EU member state nodes are circles, others squares; European nodes are black, others gray; dotted edges represent positive net ties, negative solid; edge thicknesses reflect the stronger of the two connections between nodes (this being a directed network); the size of each arrowhead reflects the estimated strength of the tie in that direction. Only ties significant at the .05 level are depicted.[32]

5. Conclusion and discussion

In [19], we estimated single-lag STAR models of ALM policy using binary contiguity (borders) weights matrices and a sample of European states over the period 1987–1998. Our estimated coefficients on the spatial lags in those regressions were negative and statistically significant, and we argued that those 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 (i.e., mechanisms) 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 to a great extent due to cross-border spillovers as we had suggested. We also

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

31 The Appendix gives analogous graphs for the 1981 and 2001 estimates. The placement algorithm for the figure is NetDraw's "Spring Embedding" with "Distances + Node Repulsion" as the layout criteria, using 100 iterations following a random initial layout with "distance between components" set to 5 and "proximity" being determined using "geodesic distances".

32 We might plot as singletons nodes with all their estimated in- and out-dependencies insignificant, or of negligible strength, but there are no such ones in this example.

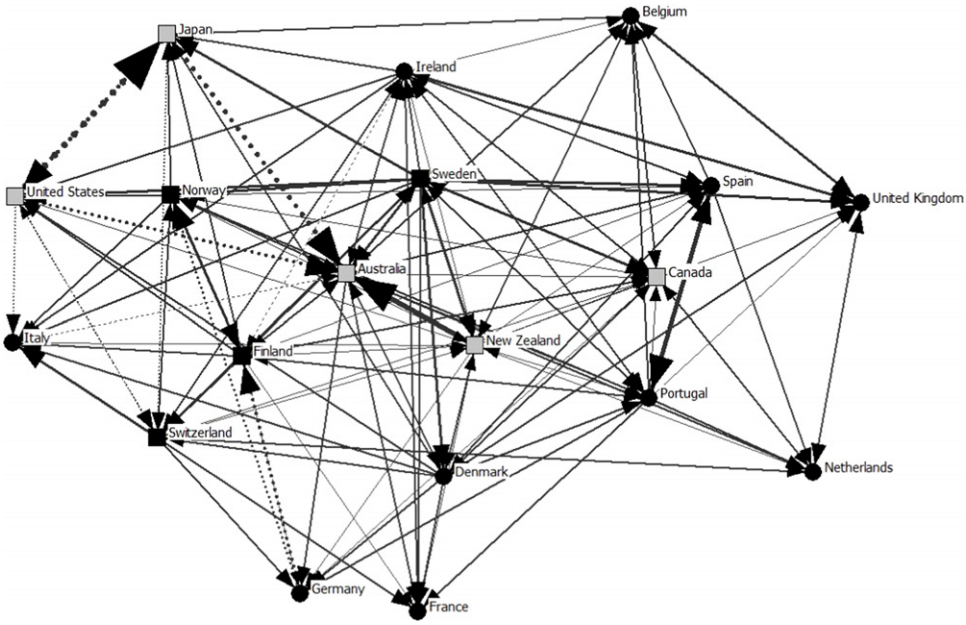

Fig. 3. The estimated network of ALM policy interdependence, 1991.

find now some evidence of positive dependence deriving from trade-related competition, supporting globalization-induced competitive races sorts of arguments, and that policymakers follow most closely foreign policymakers from similar states, where similarity is gauged by the level of the policy in question; i.e., we find evidence of *homophily* in policymaking.

Methodologically, we have proposed the m-STAR model as providing a simple way to estimate the patterns and pathways of network interdependence simultaneously with estimating the effect of those network interdependencies on units' actions. We would suggest, in particular, that the m-STAR model is well conceived to address two different kinds of aims that empirical researchers may have in spatial and spatiotemporal modeling. On the one hand, as means of simultaneous estimation of net interdependencies and the effects of those interdependencies, the m-STAR model confronts the heavy reliance of spatial and spatiotemporal empirical results on the relative and absolute accuracy and power of the researcher's specifications of the spatial ($\mathbf{W}$) and non-spatial ($\mathbf{X}$) components of the model. On the other hand, the m-STAR model also directly addresses social science researchers' interest in testing, or gauging the relative strengths of, alternative potential paths, or mechanisms, by which network *cum* spatial interdependence may operate.

Within this framework, we have suggested and started upon the more ambitious agenda of endogenizing those two co-evolutionary components: network/interdependence patterns and unit behavior/action. At present, however, our strategy for addressing the simultaneity of network selection and network/spatial contagion/interdependence effects was *the poor man's exogeneity* of merely time lagging any endogenous components of $\mathbf{W}$ and then assuming the strong conditions for the validity of that identification strategy to hold. As we emphasized, however, this does not really address true or effective simultaneity, which seems a likely situation for C & IPE contexts at least. We therefore propose in conclusion a two-step estimation procedure for future exploration. First, apply spatial instrumental variables (or two-stage least squares or the generalized method of moments: S-IV, S-2SLS, S-GMM; see, e.g., [12]) to obtain consistent (and asymptotically efficient) estimates of the endogenous w_{ij} and their estimated variance–covariance matrix. That is, use the exogenous $\mathbf{W}_r$ and $\mathbf{X}$ – recall that their exogeneity was already assumed – to instrument the ($\mathbf{y}$ in the) endogenous $\mathbf{L}$. Then, draw from this estimated multivariate distribution the ($\hat{\mathbf{y}}$ in) $\hat{\mathbf{L}}$ to insert in the appropriate likelihood, either the conditional (16) or the unconditional extension of (6). Maximize this likelihood q times, each

time with new draws from that first-stage instrumented $\hat{\mathbf{L}}$. The averages of these q second-stage S-ML estimates should then be the point estimates of parameters, and the estimated variance–covariance matrix of those parameter estimates should be the average of the estimated variance–covariance matrices across the iterations plus $(1 + q)$ times the sample variance–covariance matrix in the point estimates across iterations (as, e.g., in multiple imputation). This estimator should inherent desirable properties from S-IV (or S-2SLS or S-GMM) and S-ML as far as we can intuit, although we can offer no proof of properties yet. Assessment of the proposed estimator and direct comparison to network co-evolution approaches are the obvious next steps.

Acknowledgements

This research was supported in part by NSF grant #0318045.

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

Appendix. Supplementary data

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.stamet.2009.11.005. Replication data and code available from the authors' web pages.

References

[1] M. Allers, J.P. Elhorst, Tax mimicking and yardstick competition among local governments in the Netherlands, University of Groningen, 2005, unpublished manuscript.
[2] L. Anselin, Spatial Econometrics: Methods and Models, Kluwer Academic, Boston, 1988.
[3] L. Anselin, Spatial externalities, spatial multipliers and spatial econometrics, International Regional Science Review 26 (2) (2003) 153–166.
[4] S. Basinger, M. Hallerberg, Remodeling the competition for capital: How domestic politics erases the race to the bottom, American Political Science Review 98 (2) (2004) 261–276.
[5] N. Beck, K.S. Gleditsch, K. Beardsley, Space is more than geography: Using spatial econometrics in the study of political economy, International Studies Quarterly 50 (2006) 27–44.
[6] A. Bhattacharjee, C. Jensen-Butler, Estimation of spatial weights matrix, with an application to diffusion in housing demand, School of Economics and Finance, University of St. Andrews, UK, 2006. http://www.st-andrews.ac.uk/crieff/papers/dp0519.pdf.
[7] B. Boots, C. Dufournaud, A programming approach to minimizing and maximizing spatial autocorrelation statistics, Geographical Analysis 26 (1994) 54–66.
[8] J.K. Brueckner, L.A. Saavedra, Do local governments engage in strategic property-tax competition? National Tax Journal 54 (2) (2001) 203–229.
[9] D. Cameron, The expansion of the public economy: A comparative analysis, American Political Science Review 72 (4) (1978) 1243–1261.
[10] A.C. Case, H.S. Rosen, J.R. Hines, Budget spillovers and fiscal policy interdependence: Evidence from the states, Journal of Public Economics 52 (3) (1993) 285–307.
[11] X. Chen, T. Conley, A new semiparametric spatial model for panel time series, Journal of Econometrics 105 (1) (2001) 59–83.
[12] T. Conley, GMM estimation with cross-sectional dependence, Journal of Econometrics 92 (1) (1999) 1–45.
[13] J.P. Elhorst, Dynamic models in space and time, Geographical Analysis 33 (2001) 119–140.
[14] J.P. Elhorst, Specification and estimation of spatial panel data models, International Regional Science Review 26 (2003) 244–268.
[15] J.P. Elhorst, Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels, Geographical Analysis 37 (2005) 85–106.

[16] E. Fernandez-Vazquez, J. Rodriguez-Valez, Taking off some hoods: Estimating spatial models with a non-arbitary *W* matrix, Paper Presented at the 2007 Meeting of the Spatial Econometric Association, 2007. http://fp.paceprojects.f9.co.uk/Vazquez.pdf.
[17] R. Franzese, J. Hays, Empirical modeling strategies for spatial interdependence: Omitted-variable vs. simultaneity biases, Presented at the 21st Summer Meeting of the Society for Political Methodology, 2004.
[18] R. Franzese, J. Hays, Spatiotemporal models for political-science panel and time-series-cross-section data, Presented at the Summer Meeting of the Political Methodology Society, 2006.
[19] R. Franzese, J. Hays, Strategic interaction among EU governments in active-labor-market policymaking: Subsidiarity and policy coordination under the European employment strategy, European Union Politics 7 (2) (2006) 167–189.
[20] R. Franzese, J. Hays, Empirical models of international capital-tax competition, in: G. Gregoriou, C. Read (Eds.), International Taxation Handbook, Elsevier Press, 2007, pp. 43–72.
[21] R. Franzese, J. Hays, Spatial-econometric models of cross-sectional interdependence in political-science panel and time-series-cross-section data, Political Analysis 15 (2) (2007) 140–164.
[22] R. Franzese, J. Hays, Empirical models of spatial interdependence, in: J. Box-Steffensmeier, H Brady, D. Collier (Eds.), Oxford Handbook of Political Methodology, Oxford UP, Oxford, 2008, pp. 570–604.
[23] R. Franzese, J. Hays, Interdependence in comparative politics: Theoretical- & empirical-model specification, estimation, interpretation, and presentation, Comparative Political Studies 41 (4–5) (2008) 742–780.
[24] R. Franzese, J. Hays, Spatial Econometric Models of Interdependence, Book manuscript, 2008.
[25] R. Franzese, J. Hays, A comparison of the small-sample properties of several estimators for spatial-lag count-models, Presented at the Summer Meeting of the Political Methodology Society, 2009.
[26] R. Franzese, J. Hays, The spatial probit model of interdependent binary outcomes: Estimation, interpretation, and presentation, Presented at the Annual Meeting of the Public Choice Society, 2009.
[27] P.G. Fredriksson, D.L. Millimet, Strategic interaction and the determination of environmental policy across US states, Journal of Urban Economics 51 (1) (2002) 101–122.
[28] G. Garrett, Partisan Politics in the Global Economy, Cambridge UP, New York, 1998.
[29] A. Getis, J. Aldstadt, Constructing the spatial weights matrix using a local statistic, Geographical Analysis 36 (2) (2004) 90–104.
[30] A. Getis, J. Aldstadt, Using AMOEBA to create a spatial weights matrix and identify spatial clusters, Geographical Analysis 38 (4) (2006) 327–343.
[31] A. Getis, D. Griffith, Comparative spatial filtering in regression analysis, Geographical Analysis 34 (2002) 130–140.
[32] U. Grenander, Abstract Inference, Wiley Series, New York, 1981.
[33] D.A. Griffith, Some guidelines for specifying the geographic weight matrix contained in spatial statistical models, in: S.L. Arlinghaus (Ed.), Practical Handbook of Spatial Statistics, CRC, Boca Raton, 1996, pp. 65–82.
[34] D. Griffith, Spatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization, Springer-Verlag, Berlin, Germany, 2003.
[35] D.A. Griffith, Assessing spatial dependence in count data: Winsorized and spatial filter specification alternatives to the auto-Poisson model, Geographical Analysis 38 (2) (2006) 160–179.
[36] J. Hays, Bucking the system: Using simulation methods to estimate and analyze systems of equations with qualitative and limited dependent variables, Paper given at the St. Louis Area Methods Meeting, SLAMM, Washington University in St. Louis, 2009. http://slamm.wustl.edu.
[37] J. Hays, Globalization and the New Politics of Embedded Liberalism, Oxford UP, Oxford, 2009.
[38] J. Hays, S. Ehrlich, C. Peinhardt, Government spending and public support for trade in the OECD: An empirical test of the embedded liberalism thesis, International Organization 59 (2) (2005) 473–494.
[39] J. Hays, A. Kachi, Estimating interdependent duration models with an application to government formation and survival, Paper Given at the Political Methodology Society Summer Meeting, 2008. http://polmeth.wustl.edu/conferences/methods2008/program.html.
[40] J.J. Heckman, Dummy endogenous variables in a simultaneous equation system, Econometrica 46 (4) (1978) 931–959.
[41] P.D. Hoff, A.E. Raftery, M.S. Handcock, Latent space approaches to social network analysis, Journal of the American Statistical Association 97 (460) (2002) 1090–1099.
[42] P.D. Hoff, M.D. Ward, Modeling dependencies in international relations networks, Political Analysis 12 (2) (2004) 160–175.
[43] P.D. Hoff, A. Westveld, A bayesian mixed effects model for longitudinal social network data, University of Washington, 2007 (unpublished).
[44] T. Iversen, T.R. Cusack, The causes of welfare state expansion: Deindustrialization or globalization? World Politics 52 (3) (2000) 313–349.
[45] P. Katzenstein, Small States in World Markets, Cornell University Press, Ithaca, NY, 1985.
[46] H. Kitschelt, The Transformation of European Social Democracy, Cambridge UP, 1994.
[47] S. Kooijman, Some remarks on the statistical analysis of grids especially with respect to ecology, Annals of Systems Research 5 (1976).
[48] Leenders, T.A.J. Roger, Structure and influence, Dissertation, ICS/University of Groningen, ISBN 90-5170-329-5, 1995.
[49] Leenders, T.A.J. Roger, Longitudinal behavior of network structures and actor attributes: Modeling interdependence of contagion and selection, in: P. Doreian, F. Stokman (Eds.), Evolution of Social Networks, Routledge, Amsterdam, 1997, pp. 165–184.
[50] J.P. LeSage, R. Pace, Introduction to Spatial Econometrics, CRC Press, 2009.
[51] G. Meen, Spatial aggregation, spatial dependence and predictability in the UK housing market, Housing Studies 11 (3) (1996) 345–372.
[52] S. Openshaw, Optimal zoning systems for spatial interaction models, Environment and Planning A 9 (1977) 169–184.
[53] J.K. Ord, A. Getis, Local spatial autocorrelation statistics: Distributional issues and an application, Geographical Analysis 27 (1995) 286–306.
[54] H.G. Overman, D. Puga, Regional unemployment clusters: Nearness matters within and across Europe's borders, Economic Policy 17 (34) (2002) 117–147.

[55] J. Pinkse, M.E. Slade, C. Brett, Spatial price competition: A semiparametric approach, Econometrica 70 (3) (2002) 1111–1153.
[56] T. Plümper, E. Neumayer, Model specification in the analysis of spatial dependence, SSRN Working Paper #1092113, 2008.
[57] M. Redoano, Fiscal interactions among European countries, Warwick Economic Research Papers No. 680, 2003.
[58] D. Rodrik, Has Globalization Gone Too Far? Institute for International Economics, Washington, DC, 1997.
[59] J.G. Ruggie, International regimes, transactions, and change: Embedded liberalism in the postwar economic order, International Organization 36 (2) (1982) 195–231.
[60] T.E. Smith, Estimation bias in spatial models with strongly connected weight matrices, Geographical Analysis 41 (2009) 307–332.
[61] T.A.B. Snijders, Stochastic actor-oriented models for network change, in: P. Doreian, F. Stokman (Eds.), Evolution of Social Networks, Routledge, Amsterdam, 1997, pp. 185–208.
[62] T.A.B. Snijders, The Statistical Evaluation of Social Network Dynamics, in: M.E. Sobel, M.P. Becker (Eds.), Sociological Methodology, vol. 31, 2001, pp. 361–395.
[63] T.A.B. Snijders, Models for longitudinal network data, in: P. Carrington, J. Scott, S. Wasserman (Eds.), Models and Methods in Social Network Analysis, Cambridge UP, New York, 2005 (Chapter 11).
[64] T.A.B. Snijders, S.P. Borgatti, Non-parametric standard errors and tests for network statistics, Connections 22 (2) (1999) 161–170.
[65] T.A.B. Snijders, C. Steglich, P. West, Applying SIENA: An illustrative analysis of the co-evolution of adolescents' friendship networks, taste in music, and alcohol consumption, 2006.
[66] T.A.B. Snijders, C. Steglich, M. Schweinberger, Modeling the co-evolution of networks and behavior, in: K. van Montfort, H. Oud, A. Satorra (Eds.), Longitudinal Models in the Behavioral and Related Sciences, Lawrence Erlbaum, Mahwah, NJ, 2007, pp. 41–71 (Chapter 3).
[67] D. Swank, Global Capital, Political Institutions, and Policy Change in Developed Welfare States, Cambridge UP, Cambridge, 2002.