Spatial-Econometric Models of Interdependence

(*A Book Prospectus*)

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

Jude C. Hays (jude.hays@gmail.com)
Associate Professor of Political Science
The University of Pittsburgh

25 August 2010

ABSTRACT

This book covers the specification, estimation, interpretation, and presentation of empirical models that reflect spatial and spatiotemporal interdependence among observations.

Chapter 1 provides overview of the substance and theory of interdependence in the social sciences, with an emphasis on political science and, especially therein, comparative and international politics and political economy. Part A surveys the broad substantive range in which, and the multifarious mechanisms by which, spatial interdependence arises across these subjects. Part B first presents and discusses a general theoretical model which demonstrates that interdependence arises whenever the marginal utility to unit *i* of its actions/policies depends on the actions/policies of some unit(s) *j*. The model illuminates the logical connections from positive externalities to strategic-substitute relations, free-rider incentives, first-/early-mover advantages, and strategic-delay/war-of-attrition dynamics on the one hand, and from negative externalities to strategic-complement relations, beggar-thy-neighbor incentives, last-/late-mover advantages, and competitive-race (e.g., *race-to-the-bottom*) dynamics on the other. It then presents, and derives empirically testable propositions from, a more-specific theoretical model in the substantive context of international capital-tax competition for foreign investment before offering briefer overview of theoretical models of interdependence in other substantive areas across the social sciences.

Chapter 2 introduces and explores the serious challenges for empirical analysis posed by interdependence. It first contrasts two of three possible sources of spatial correlation—correlated unit-specific and/or exogenous-external explanators (which we call *common exposure*) and/or true interdependence or *contagion*—distinguishing between which is the essence of *Galton's* (famous) *Problem*. Whether these spatial associations or true interdependencies arise in stochastic and/or systematic components creates a thorny additional complication of considerable importance because each implies substantively very different spatial dynamics. It then discusses in this vein several empirical observations commonly thought to be evidence of interdependence—e.g., S-shaped patterns of policy-adoption or temporal fixed-effects or regional dummies—and explains how they, unlike well-specified spatial-econometric models with spatial lags, are incapable of distinguishing among possible sources of spatial correlation. Next it introduces and discusses various measures and tests of spatial correlation and/or interdependence, some of which are capable of offering at least some judgment on their alterative possible sources. Finally, the third possible source of observed spatial association—*selection* by (i.e., endogeneity to) the units of observation of the dimension along which proximity is measured—is introduced and discussed.

Chapter 3 turns to the technical specification and estimation of empirical models of interdependence. It begins by presenting a generic empirical model with spatial lag, unit-specific and exogenous-external shocks, and the interaction of the latter two. Such a model corresponds to modern theoretical and substantive expectations in political science and highlights the empirical difficulty of distinguishing the possible sources of spatial correlation. It then introduces and discusses four common estimation strategies for such spatial-lag models: non-spatial OLS (or ML), OLS (or ML) with spatial lags, spatial method-of-moments (IV/2SLS/GMM), and fully specified spatial ML. It next identifies four broad classes of models with spatial interdependence—spatial-error, spatial-lag, spatiotemporal-lag models, and spatial- and spatiotemporal-lag qualitative-dependent-variable models. The final section discusses more technically than did Chapter 2 the challenges in estimating these sorts of models for these estimators: omitted-variable and simultaneity biases for the two simpler estimators, specification and measurement error for all of them, and the many serious issues surrounding the all-important spatial-connectivity matrix, **W**, that describes the relative strengths of interdependence from units *j* to *i*. This last part includes discussion of (1) the pre-specification of **W**, which is current practice and is perhaps the most crucial stage of empirical spatial analysis (*inter alia*, this part explains and explores *Achen's Problem* with lagged-dependent-

variable models, namely that insufficiently accurate dynamic specifications push the LDV toward "stealing explanatory power" from the other regressors, in the spatial context), (2) how estimators' properties tend to vary, often in complicated fashion, with the nature of **W** (this necessitates the many repeated sets of simulations in Chapter 4), (3) issues raised by multiple-**W** models (used in the literature, for example, to evaluate the strength of several alternative mechanisms of interdependence that may be operating), and (4) some inroads into the parameterization and estimation of **W**, including the complex case – in both the continuous-interval and the qualitative-limited dependent-variable contexts – of endogenous **W**, in which the pattern or network of connectivity between units and the behavior of the units *coevolve*, i.e., are jointly endogenously determined.

Chapter 4 considers the properties of these estimators. It first derives analytically (in simple cases, asymptotically) the respective omitted-variable and simultaneity biases of the non-spatial estimators and the LS or ML estimators for models with spatial lags, spatial and temporal lags, and multiple spatial lags. It then evaluates by simulation the bias, efficiency, and standard-error-accuracy performance of all the estimators in richer contexts and limited samples. These simulations explore, first, data-generating processes and empirical specifications matching the generic empirical model of Chapter 3, with common exogenous-external shocks and uniform interdependence among spatial units. The spatial lag is thus the simple average of the dependent variable in other units, representing a "rough & ready" spatial-lag proxy to which empirical researchers may often resort. *Galton's Problem* manifests severely here, given the close similarity of uniform interdependence and common exogenous-external shocks. Across simulations, the estimators are evaluated first with empirical specifications matching the true data-generating process; then with measurement-*cum*-specification error in the empirical model of the exogenous shock or the spatial-interdependence pattern; and then, for the moment (instrumentation) estimators, with imperfectly exogenous instruments under varying degrees of spatial correlation and endogeneity among right- and left-hand-side variables and, for the fully specified ML estimators, with varyingly sizable violations of their strong distributional assumptions. For the spatiotemporal models, the simulations also explore *Achen's Problem* with lagged-dependent-variable (LDV) models in the spatial context. All these simulation exercises are then repeated for several realistically non-homogenous patterns of spatial-interdependence literally drawn from substantive literatures; i.e., elements of the spatial-connectivity matrix used for experimental consideration are drawn directly from actual datasets used in the literature. One set of patterns reflects international economic relations among the units; another reflects "co-memberships" (e.g., in organizations, language or cultural groups, etc.; shared borders, dyads, etc.) among the units; a third reflects the geographic or Euclidean distance between the units; and the fourth pattern reflects strategic-complement (cooperative) and -substitute (competitive) relations between the units. The last two sections of the chapter conduct similar sets of simulations of multiple-**W** and spatial-lag qualitative-dependent-variable models.

Chapter 5 shows how to calculate, and offers suggestions on presenting (via tabular feedback grids, spatiotemporal response-path graphs, and maps), estimates and standard errors (by the delta method and by simulation) of estimates of spatial and spatiotemporal multipliers, effects, dynamics (response paths), and long-run steady-state effects.

Chapter 6 illustrates all this via replications and extensions of existing work using spatial lags or failing to use them when theory and substance clearly imply them (Swank & Steinmo 2002, Hays 2003, Basinger and Hallerberg 2004, Beck, Gleditsch, & Beardsley 2006, Volden 2006, Franzese & Hays 2006).

Two appendices provide Stata™, MatLab, and/or R code to implement all specification, estimation, testing, and presentation procedures described in the text, and all data and code to replicate the empirical work.

OUTLINE / TABLE OF CONTENTS

I. **(CHAPTER 1)** The Substance & Theory of Spatial Interdependence

A. The Broad Substantive Range of Spatial Interdependence

1. Topics and Literature Review

2. Mechanisms: Coercion; Competition, Cooperation, and Externalities; Learning and Emulation; Migration

3. => Tobler's First Law of Geography (*all things related, near things more so*), with Beck, Gleditsch, & Beardsley's addendum (*space is more than geography*), and Franzese & Hays' corollary (*space is substance, not nuisance*).

B. Theoretical Models Strategic Interdependence:

1. A General Theoretical Model of Interdependence

a) Basic Principle: if marginal utility of *i*'s actions depend on *j*'s actions, then interdependence

b) General Properties:

(1) Positive Externalities => Strategic Substitutes wherein policies/actions move in opposite directions, i.e., negative interdependence, in which *free-rider* problems & late-mover advantages, & so strategic delay/wars of attrition dynamics, arise.

(2) Negative Externalities => Strategic Complements wherein policies/actions move in same directions, i.e., positive interdependence, in which *beggar-thy-neighbor* problems & early-mover advantages, & so *races-to-the-bottom* (or *top*, or, generally, competitive races), arise.

c) Examples & Discussion

2. Some Specific Theoretical Models of Interdependence

a) The Political Economy of Globalization, Capital-Tax Competition, and Domestic Fiscal-Policy Autonomy

(1) Review of theories in previous literature; all imply interdependence, but most previous empirical specifications fail to incorporate that implication.

(2) A citizen-candidate model of capital-tax competition and foreign investment (from Persson & Tabellini 2000).

(a) Also implies contagion & renders it formally explicit.

(b) Discuss model & examples of derived comparative statics

b) Briefer overview of several other specific theoretical models, in other contexts within and beyond comparative/international political economy.

II. **(CHAPTER 2)** The Empirical Challenge of Spatial Interdependence

A. Galton's Problem: 3 Sources of Spatial Correlation, and a modern complication

1. *Common exposure*: i.e., correlated unit-specific &/or exogenous-external

conditions/explanators (examples)

2. *Contagion*: a.k.a., true interdependence (examples)

a) complication: distinguish error/stochastic-term correlation or interdependence from systematic-component/outcome interdependence: implied spatial-dynamics differ substantively importantly

3. *Galton's Problem* is that these things look very much alike empirically, and so the fact that any, all, or any combination may operate in any given context presents an extreme empirical challenge to distinguish them (by any estimation strategy or, for that matter, by any empirical methodology, quantitative or qualitative).

4. *Selection*: units may select or influence the pattern of connections between them; yields *Galton's Problem plus* (examples)

B. Empirical observations often taken as evidence of interdependence that not necessarily

1. Dummies for Groups/Regions

2. S-Shaped Adoption Curves

3. Measures & Tests of Spatial Correlation/Clustering (briefly introduced)

4. Parks FGLS, PCSE, & other estimation strategies based on error-covariance matrices with spatial patterns

C. Measures and tests of spatial correlation and interdependence:

1. Discusses how each of the above by itself and, for that matter, a spatial-lag model by itself fail to distinguish among the multiple possible sources of spatial correlation, including interdependence, and so are not necessarily evidence of the latter.

2. Elaborates some appropriate nested & non-nested test-strategies of fuller models

III. **(CHAPTER 3)** Specifying & Estimating Empirical Models of Spatial & Spatiotemporal Interdependence

A. A Generic Empirical Model of Open Socio-Political-Economies with Context-Conditionality and Interdependence

B. Estimators for Spatial-Interdependence Models

1. Non-Spatial Least-Squares and Maximum-Likelihood models (inconsistent: omitted-variable bias)

2. Least-Squares and (Quasi) Maximum-Likelihood with Spatial-Lag regressors (inconsistent: simultaneity bias)

3. Spatial Method-of-Moments Estimators:

a) Spatial Instrumental-Variables / Two-Stage-Least-Squares (consistent but inefficient)

b) Spatial Generalized-Method-of-Moments (consistent and asymptotically efficient)

4. Spatial Maximum-Likelihood (FIML) (consistent and asymptotically efficient and normal)

C. Classes of Spatial-Econometric Models

1. Spatial-Error Models (spatial interdependence in the stochastic component)
    a) Utility & Limits of Consistent Variance-Covariance Estimation (PCSE's)
    b) A Simple, Data-Driven Efficiency-Enhancement of PCSE: Utility, Limits
    c) True Spatial-Error Models
2. Spatial-Lag Models (spatial interdependence in systematic component or *y*)
    a) Can address *Galton's Problem*; key is the relative & absolute accuracy and power of the spatial and non-spatial aspects of the empirical model.
    b) As in time-series, modern view is increasingly that this is way to proceed.
3. Spatiotemporal-Lag Models
    a) Time-lagged spatial interdependence: utility and limits. Spatial LS or ML consistent under some conditions. Give and discuss those conditions.
    b) Contemporaneous spatial interdependence
4. Spatial- and Spatiotemporal-Lag Qualitative-Dependent-Variable Models:
    a) Spatial Probit
    b) Spatial Duration Models
    c) Spatial Count Models
5. Estimated-**W** and Endogenous-**W** (Coevolution) Models:
    a) Continuous-Interval Dependent-Variable & **W** Models
    b) Qualitative Dependent-Variable & **W** Models

D. Challenges for Estimation and Inference
1. Simultaneity of the simple estimation strategies (LS/ML, S-OLS/Quasi-ML)
2. Specification & Measurement Error
    a) In the non-spatial components of the model
    b) In the spatial-connectivity matrix
    c) In instruments' exogeneity or in likelihood
3. **W**: The Spatial-Connectivity Matrix of Relative Interdependence
    a) **W** commonly pre-specified; crucial stage of spatial analysis
        (1) Examples given, discussion
        (2) Achen's Problem with lagged-dependent-variable (LDV) models
            (a) Review: Time-LDV v. AR(p) residuals
            (b) Analogous issue in spatial-LDV v. spatial-error models
    b) Estimators' properties tend to vary, often in complicated fashion, with **W**
    c) Multiple-**W** Models: raise issues beyond the usual ones associated with controls (e.g., interpretation, colinearity) due to simultaneity and "overlap"
    d) Inroads into parameterization and estimation of **W**

IV. **(CHAPTER 4)** Properties and Performance of Spatial Estimators

A. Analytic Results

1. In 2-unit, 2-regressor case: omitted-variable bias (OVB) in non-spatial OLS and simultaneity in Spatial OLS

2. General Derivation of Simultaneity Biases in Spatial and Spatiotemporal Model

3. Simultaneity Biases in the multiple-**W** case.

B. Simulation Results: using generic empirical model of interdependence, and:

1. Interdependence Processes/Patterns of Interdependence:

a) Uniform/Homogenous Interdependence: The full set of experiments described are conducted first with this specification; then repeated for all of the following, on which the reports will be briefer.

b) Interdependence proportional to some economic/strategic connection; w_{ij} drawn from actual datasets in lit: trade/capital flows, trade/capital mobility,

c) Strategic complements/substitutes (cooperation/competition) relations between units; w_{ij} drawn from uniform-distribution {-1…1}.

d) Unit interdependence due to "co-membership" (borders, organizations, culturo-linguistic-religious groups, dyads, etc.); w_{ij} drawn from actual datasets in literature. Considerations in exploring this parameter space:

(1) Overlapping v. non-overlapping group-memberships.

(2) Complete group membership (no isolates)

(3) Size & number of these groups relative to totals

e) Unit interdependence proportional to geographic or Euclidean distance b/w units (drawn from actual datasets in the literature)

2. Specification and Measurement Error:

a) Fully accurate specification and measurement: The full set of experiments are conducted first thusly; then repeated for all of the following.

b) Measurement/specification error in exogenous-external factors

c) Measurement/specification error in **W**

d) Exogenous unit-specific and exogenous-external explanators vs. varying magnitudes of endogeneity from left- to right-hand-side as well as varying magnitudes spatial interdependence among left- and right-hand-side.

3. Spatiotemporal Lag Models

4. Multiple-**W** Models: Parameterized, Estimated, & Endogenous **W**

5. Spatial- and Spatiotemporal-Lag Models of Qualitative Dependent-Variables

6. Systems of Spatial Endogenous Equations: Two Kinds of Simultaneity, spatial across units and endogenous regressors within units.

V. **(CHAPTER 5)** Interpretation & Presentation of Spatial & Spatiotemporal Effects & Dynamics:

A. Calculating Spatial and Spatiotemporal Multipliers, Effects, and Dynamics

1. The Spatial Multiplier and Contemporaneous Spatial Effects
2. The Spatiotemporal Multiplier and Spatiotemporal Response-Paths and Long-Run Spatiotemporal Steady-States
3. Spatial & Spatiotemporal Effects & Dynamics in Qualitative Dependent-Variable Models
4. Spatial & Spatiotemporal Effects & Dynamics in Coevolutionary (Endogenous-**W**) Models

B. Calculating Standard Errors for S and ST Multipliers, Effects, and Dynamics
   1. Delta Method
   2. Simulation Methods

C. Presenting Estimates of Spatial and Spatiotemporal Effects and Dynamics
   1. Spatial and Spatiotemporal Effect-Grids
   2. Plotting Spatiotemporal Response-Paths
   3. Maps (and Dynamic Maps) of Spatial and Spatiotemporal Effects

VI. **(CHAPTER 6)** Empirical Applications: Replications & extensions (of existing work using spatial lags or failing to use them when theory & substance imply them, ordered by mechanism of interdependence)

A. Competition: Globalization, Capital-Tax Competition for Investment, & Domestic-Policy Autonomy (Swank&Steinmo APSR '02; Hays WP '03; Basinger&Hallerberg APSR '04)

B. Learning & Emulation: Volden (AJPS 2006) on US State Adoption of Policy Innovations

C. Competition and Negative Externalities: Beck et al. (ISQ 2006) on International Conflict and Trade

D. Cooperation & Positive Externalities: Franzese & Hays (EUP 2006) on ALMP in the EU

E. Strategic: World War I entry (Franzese et al. 2010, Hays 2009, Hays & Kachi 2009)

F. Multiple Mechanisms: Extended ALMP application (Hays et al. 2010)

VII. (Web) Appendices:

A. Stata™ code to implement all of the procedures described in the text.

B. Data and code to replicate all of the estimation results given in the text.