Hays, J.C., Kachi, A., Franzese, R.J.

7 October 2009

“A spatial model incorporating dynamic, endogenous network interdependence: a political science application”

WEB APPENDIX

CONTENTS:

Pp. 1-3: A topically organized reference list for applied spatial-econometric modeling in social science, emphasizing political science.
P. 4: A brief description of the intellectual-historical genesis of Galton’s Problem.
Pp. 8-10: Analytic and simulation results demonstrating that the S-ML estimator described and applied in the text perform well and greatly outperform simpler least-squares estimators.
P. 11: Additional network-estimates graphs from the illustrative empirical analysis.
Pp. 12-33: Complete citation list for the paper and the appendix.

Expanded References

(The list below follows the order of the topics mentioned in the introduction of Franzese, Hays, & Kachi 2009. The complete reference list appears at the end of this Appendix.)

A Subject-Organized Reference List for Applied Spatial-Modeling in Political Science:


EMPirical ATTention to the InHERent INTERDEPENDENCE OF INTERnaTional


ON INTERDEPENDENCE OF LEGISLATORS’ VOTES (MODELED SPATIALLY): See, for example, Lacombe & Shaughnessy 2005.


ON INTERDEPENDENCE OF CANDIDATE QUALITIES, CONTRIBUTIONS, OR STRATEGIES: See, for example, Goldenberg et al. 1986; Mizruchi 1989; Krasno et al. 1994; Cho 2003; Gimpel et al. 2006.

FOR SPATIAL MODELS OF THE INTERDEPENDENCE OF THE PROBABILITIES AND OUTCOMES OF COUPS: e.g., Li & Thompson 1975; OF RIOTS: e.g., Govea & West 1981; OF CIVIL WARS: e.g., Murdoch & Sandler 2004, Buehaug & Rød 2006; OF REVOLUTIONS: e.g., Brinks & Coggede 2006.

ON INTERDEPENDENCE IN TREATY SIGNING: see, e.g., Murdoch et al. 2003.

ON INTERDEPENDENCE IN TERRORIST ORIGINS AND TARGETS: see, e.g., Brathwaite & Li 2008.


ON INTERDEPENDENCE IN VIOLENCE AND CRIME: see, e.g., Grattet et al. 1998; Myers


ON INTERDEPENDENCE IN MACROECONOMIC PERFORMANCE: see, e.g., Fingleton 2003; Novo 2003; Kosfeld & Lauridsen 2004; Maza & Villaverde 2004; Kelejian et al. 2006; Mencken et al. 2006.

ON INTERDEPENDENCE IN TECHNOLOGY, MARKETING, AND OTHER FIRM STRATEGIES: see, e.g.; Abramson & Rosenkopf 1993; Geroski 2000; Strang & Macy 2001; Holloway 2002; Bradlow 2005; Autant-Berard 2006; Mizruchi et al. 2006.

ON INTERDEPENDENCE IN FERTILITY, BIRTHWEIGHT, CHILD DEVELOPMENT, OR CHILD POVERTY: see, e.g., Tolnay 1995, Montgomery & Casterline 1996; Morenoff 2003; Sampson et al. 1999; Voss et al. 2006.

The intellectual-historical genesis of Galton’s Problem:

Galton originally raised the issue thus: “[F]ull information should be given as to the degree in which the customs of the tribes and races which are compared together are independent. It might be that some of the tribes had derived them from a common source, so that they were duplicate copies of the same original. ...It would give a useful idea of the distribution of the several customs and of their relative prevalence in the world, if a map were so marked by shadings and colour as to present a picture of their geographical ranges” (Galton 1889, as quoted in Darmofal 2007). We find further historical context in http://en.wikipedia.org/wiki/Galton%27s_problem: “In [1888], Galton was present when Sir Edward Tylor presented a paper at the Royal Anthropological Institute. Tylor had compiled information on institutions of marriage and descent for 350 cultures and examined the correlations between these institutions and measures of societal complexity. Tylor interpreted his results as indications of a general evolutionary sequence, in which institutions change focus from maternal to paternal lines as societies grow more complex. Galton disagreed, noting that similarity between cultures could be due to borrowing, could be due to common descent, or could be due to evolutionary development; he maintained that without controlling for borrowing and common descent one cannot make valid inferences regarding evolutionary development. Galton’s critique has become the eponymous Galton’s Problem (Stocking 1968:175), as named by Raoul Naroll (1961, 1965), who proposed [some of] the first statistical solutions.”
Spatial-econometric and spatial-statistical approaches to spatial analysis:

Methodologically, two approaches to spatial analysis can be discerned: spatial statistics and spatial econometrics (Anselin 2002, 2006). The distinction rests, on the one hand, on the relative emphasis in spatial-econometric approaches to theoretical models of interdependence processes wherein space may often have broad meaning, well beyond geography and geometry to encompass all manner of social, economic, or political connection that induces effects from outcomes in some units on outcomes in others (Brueckner 2003; Beck et al. 2006). The spatial-lag regression model plays a starring role in that tradition (Hordijk 1974; Paelinck and Klaassen 1979; Anselin 1980, 1988, 1992; Haining 1990; LeSage 1999). According to Anselin, theory driven models deal with substantive spatial correlation (2002). This approach has its own method to model specification and estimation. The importance of spatial interdependence is tested using Wald tests and the unrestricted spatial lag model. On the other hand, spatial-error models, analysis of spatial-correlation patterns, spatial kriging, and spatial smoothing, e.g., characterize the more-exclusively data-driven approaches and the typically narrower conception of space in solely geographic/geometric terms in the longer spatial-statistics tradition (initially inspired by Sir Galton’s famous comments at the 1888 meetings of the Royal Anthropological Society, and reaching crucial methodological milestones in Whittle 1954; Cliff & Ord 1973, 1981; Besag 1974; Ord 1975; Ripley 1981; Cressie 1993). According to Anselin, this kind of modeling is driven by data problems such as measurement error. The spatial correlation is viewed as a nuisance. This approach has its own method to model specification and estimation. The presence of spatial effects is tested using Lagrange multiplier tests and the restricted non-spatial lag model.

Anselin (2002, 2006), Griffith & Paelinck (2007), and many others discern two approaches to spatial analysis: spatial statistics and spatial econometrics. “In practice, spatial econometrics and spatial statistics reflect traditions of their parent disciplines. In other words, they share much in common, with some notably differing emphases… Econometrics…is the ‘setting up of mathematical models describing economic relationships…, testing the validity of such hypotheses and estimating the parameters in order to obtain a measure of the strengths of the influences of the different independent variables’ (Bannock et al. 2003)… ‘Statistics is the science of gaining information from numerical data’ (Moore 1995, p. 2). It provides data-interrogative tools and conceptual
frameworks for gaining understanding through empirical-based induction, and involves
data acquisition, data analysis, and statistical inference” (Griffith & Paelinck 2007:210).
Thus, the distinction rests, on the one hand, on the relative emphasis in spatial-
econometric deductive approaches to theoretical models of interdependence processes,
wherein space may often have broad meaning, well beyond physical distance to
encompass all manner of social, economic, or political connection that induces effects
from outcomes in some units on outcomes in others (see, e.g., recently, from economics
and political science: Brueckner 2003; Beck et al. 2006). Spatial-lag models play starring
roles in that tradition (Hordijk 1974; Paelinck & Klaassen 1979; Anselin 1980, 1988,
1992; LeSage 1999). According to Anselin (2002), such theory-driven models deal with
substantive spatial correlation, and this approach lends itself to methods of model
specification, estimation, and evaluation that begin with the unrestricted spatial lag
model and use Wald-style testing to gauge the importance of spatial interdependence.
On the other hand, spatial-error models, analysis of spatial-correlation patterns, spatial
kriging, and spatial filtering, for examples, characterize the more-exclusively data-driven,
i.e., inductive, approaches and the more-typically narrower conception of space in solely
geographic/geometric terms in the longer spatial-statistics tradition (initially inspired by
Sir Galton’s famous comments at the 1888 meetings of the Royal Anthropological
Society, and reaching methodological milestones in Whittle 1954; Cliff & Ord 1973,
1981; Besag 1974; Ord 1975; Ripley 1981; Cressie 1993). Again following Anselin (2002),
this modeling approach stresses spatial patterns and clustering, seen more often as
driven by data problems such as measurement error, and so tends to treat spatial
correlation more as nuisance than substance. Such emphases lend themselves more
naturally to the opposite direction in model specification, estimation, and evaluation,
beginning with the restricted non-spatial model and adopting Lagrange-multiplier-style
testing.¹ We would mention additionally only how the core questions asked and sorts of
answers sought tend to differ across approaches. The core substantive distinction drawn
in our proposal between spatial association arising from common exposure, contagion, or
selection, and the challenges of separately identifying and estimating them are typically
more centrally stressed from spatial-econometric perspectives (model specified

¹ A debate chronicled in Regional Science and Urban Economics (Florax et al. 2003, 2006; Hendry 2006;
Mur & Angulo 2009) continues as to whether one can demonstrate inferential dominance of the bottom-up
or top-down strategy.
deductively for purposes of inference about processes) than from spatial-statistical ones. Indeed, the fact of association will often suffice from a spatial-statistical vantage (inductive process of gaining information from the data), so the modeling approach often makes no attempt to distinguish them. Likewise, the kinds of counterfactuals of interest to us arise more naturally in a spatial-econometric than a spatial-statistical framework.

Network analysis, finally, considers a related (but not identical) set of substantive questions regarding the structure of ties (edges, arcs, connections) between units (nodes, actors, vertices). Central questions concern characterizing (measuring) the structure of networks, explaining their genesis, and, less centrally, at least until very recently, considering the effects of network structure and of units’ location within it on units’ actions (behavior, attributes). Again until perhaps recently, questions of how other units’ (alters’) actions affect each unit (ego) via the connections given by the network seem to have been less central still. Methods for network analysis seem to have originated primarily in sciences (physics, biology, computer-science) and mathematics, rather than in statistics or econometrics, and methodological development, including the eventual importation of statistical concepts, theories, and methods, has been largely separate from either spatial statistics or spatial econometrics (again, until recently perhaps), notwithstanding the similarity of their substantive interests. The typical core questions of network analysis, thus, are subtly distinct from those of spatial econometrics and spatial statistics. Our own emphasis on Galton’s Problem, the distinguishing of common exposure from interdependence (a.k.a., contagion) sources of spatial association, takes an important additional concern from a core network-analysis question of explaining network genesis, i.e., selection (a.k.a., homophily).

---

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

Before illustrating the estimation, testing, and interpretation of our m-STAR model of network-behavior coevolution, we will demonstrate that, in fact, the S-ML estimators just described are needed and outperform simpler least-squares estimators. Analytically, the omitted-variable biases of the blind OLS estimator remain as before: $\mathbf{F}$. The simultaneity asymptotic biases (inconsistencies) of the naïve S-OLS estimator, which simply inserts the multiple spatial lags into least-squares regression, are also analogous to Error! Reference source not found., as follows:

Let $\mathbf{y} = Z\mathbf{p} + \mathbf{\varepsilon}$; where $Z = \begin{bmatrix} \mathbf{W}_i \mathbf{y} \end{bmatrix}$ and $\mathbf{p} = \begin{bmatrix} \rho_1 & \rho_2 \end{bmatrix}^T$.

Then: $\text{plim } \hat{\mathbf{p}} = \mathbf{p} + \text{plim } \left( \frac{\hat{\mathbf{Z}}\hat{\mathbf{\varepsilon}}}{n} \right)$, that is:

$$\text{plim } \hat{\mathbf{p}} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} + \frac{V(\mathbf{W}_i \mathbf{y}) \times V(\mathbf{W}_i \mathbf{y}) - C(\mathbf{W}_i \mathbf{y}, \mathbf{W}_i \mathbf{y})^T}{V(\mathbf{W}_i \mathbf{y}) \times V(\mathbf{W}_i \mathbf{y}) - C(\mathbf{W}_i \mathbf{y}, \mathbf{W}_i \mathbf{y})^T} \cdot \frac{-C(\mathbf{W}_i \mathbf{y}, \mathbf{W}_i \mathbf{y})}{V(\mathbf{W}_i \mathbf{y}) \times V(\mathbf{W}_i \mathbf{y}) - C(\mathbf{W}_i \mathbf{y}, \mathbf{W}_i \mathbf{y})^T} \cdot \frac{C(\mathbf{i}, \mathbf{W}_i \mathbf{y})}{C(\mathbf{i}, \mathbf{W}_i \mathbf{y})},$$

(1).

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

To demonstrate that an estimator is inconsistent, however, does not demonstrate that these asymptotic biases are practically large or that they outweigh other potential deficiencies of consistent estimators. Accordingly, we conduct some simple Monte Carlo simulations to explore the small-sample performance of these estimators, specifically the magnitudes of their biases, inefficiency, and standard-error inaccuracy. With the analytical results in (1), Table 1’s simulation results are easily interpreted. The covariance of the queen spatial-lag (all eight adjacent squares on a grid) and $\mathbf{\varepsilon}$ is about half that of the rook lag (only the four horizontally and vertically adjacent) and $\mathbf{\varepsilon}$. With row-standardization, eight ties, and fixed $\rho$, the strength of the interdependence/endogeneity is more diluted in queen-lag $\mathbf{W}_i \mathbf{y}$. Consequently, $\hat{\rho}_1$ on rook-lag $\mathbf{W}_i \mathbf{y}$ is overestimated, and $\hat{\rho}_2$ on queen-lag $\mathbf{W}_i \mathbf{y}$ and $\hat{\beta}_0$ on $x_0$ (which is
especially correlated with $W_1y^3$ are (badly) underestimated. The S-ML estimator also dominates impressively in efficiency and standard-error accuracy, especially for those three coefficients.

| Table 1. Estimator Comparison for m-STAR Model: S-OLS vs. S-ML |
|------------------------|--------|--------|--------|--------|
|                        | $\beta_0=1$ | $\beta_1=1$ | $\rho_1=.3$ | $\rho_2=.3$ |
| **Average Estimate**   | 0.38 / 0.24 | 0.96 / 0.97 | 0.47 / 0.47 | 0.27 / 0.29 |
| **Root Mean-Squared-Error** | 0.71 / 0.80 | 0.08 / 0.06 | 0.25 / 0.22 | 0.21 / 0.16 |
| **Average Std-Err Estimate** | 0.37 / 0.28 | 0.06 / 0.05 | 0.15 / 0.11 | 0.17 / 0.13 |
| **Overconfidence**     | 0.92 / 0.87 | 1.06 / 1.02 | 1.29 / 1.35 | 1.21 / 1.27 |

| **Average Estimate**   | 1.09 / 1.00 | 1.00 / 1.00 | 0.31 / 0.31 | 0.27 / 0.28 |
| **Root Mean-Squared-Error** | 0.34 / 0.24 | 0.07 / 0.05 | 0.12 / 0.09 | 0.14 / 0.11 |
| **Average Std-Err Estimate** | 0.31 / 0.23 | 0.06 / 0.05 | 0.12 / 0.09 | 0.14 / 0.11 |
| **Overconfidence**     | 1.05 / 1.05 | 1.03 / 1.00 | 0.98 / 1.01 | 0.98 / 1.01 |

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

Table 2 similarly evaluates the blind OLS, naive S-OLS, and S-ML estimators for our m-STAR coevolution model. S-ML again outperforms the inconsistent OLS alternatives. Notice that, with $x$ drawn independently, appreciable correlation of the regressors with the spatial lags concentrates in the time-lag; thus, omitted-variable biases of blind OLS are not severe for $\beta$ and concentrate at a noticeable 20% in $\phi$. Notice also that with endogenous $L_y$ being time-lagged as estimated and in truth, and $L$ being $|y_i-y_j|$ and so not terribly linearly related to $y$, the simultaneity biases of S-OLS concentrate in $\rho$ at a sizable +33%, whereas $x$ being drawn independently, little induced bias appears in $\beta$.

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

---

3 The simulations drew $\epsilon$ from $N(1,1)$, making the nonzero aspect of $ZWy$ concentrate in the constant, $x_\epsilon$. 
Table 2. Estimator Comparison for m-STAR Coevolution Model: OLS v. S-OLS v. S-ML

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Result</th>
<th>φ=.3</th>
<th>β=1</th>
<th>ρ=.3</th>
<th>γ=.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>Average Estimate</td>
<td>.36</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.07</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.09</td>
<td>.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.03</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>2.07</td>
<td>.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-OLS</td>
<td>Average Estimate</td>
<td>.28</td>
<td>.99</td>
<td>.41</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.03</td>
<td>.05</td>
<td>.06</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.04</td>
<td>.05</td>
<td>.129</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.03</td>
<td>.04</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>1.02</td>
<td>1.05</td>
<td>1.19</td>
<td>1.01</td>
</tr>
<tr>
<td>S-ML</td>
<td>Average Estimate</td>
<td>.29</td>
<td>1.00</td>
<td>.31</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>.03</td>
<td>.05</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Root Mean-Squared-Error</td>
<td>.03</td>
<td>.05</td>
<td>.05</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>Average Std-Err Estimate</td>
<td>.03</td>
<td>.04</td>
<td>.04</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>1.02</td>
<td>1.04</td>
<td>1.11</td>
<td>1.10</td>
</tr>
</tbody>
</table>

Monte Carlo (1000 Trials) Results for $y_t = \rho Wy_t + \gamma Ly_{t-1} + \phi y_{t-1} + \beta x_t + \epsilon_t$, with $W=48$ contiguous US-state adjacency pattern (row-stdzd); $\rho=.3, \gamma=.3, \phi=.3, \beta=1$; and $N=48, T=10$.

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

**Figure 4:** The Estimated Network of ALM-policy Interdependence, 1981

**Figure 6:** The Estimated Network of ALM-policy Interdependence, 2001


45.


Freeman, R. Topel, and B Swedenborg, Eds. Chicago: University of Chicago.


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


Holmes, T.J. 2006. “Geographic Spillover of Unionism.” Federal Reserve Bank of Minneapolis Research Department Staff Report 368.


Taiwan: Spatial regression with disjoint neighborhoods,” Political Research Quarterly 59:35–46.


