Time-Series-Cross-Section Data Analysis

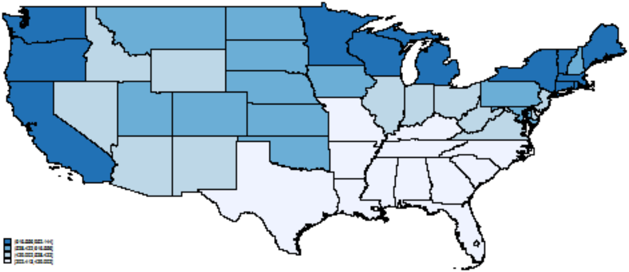
Diagnosing and Modeling Spatial Dependence

Jude C. Hays
jch61@pitt.edu

March 3, 2021

Outline

Figure: Weekly AFDC Benefits



Why do welfare benefits cluster geographically?

Interdependence vs. Clustering

Why do welfare benefits and regime types cluster geographically?

- Interdependence: Welfare migration induces a localized race-to-the-bottom in benefits. Diffusion in regime type. E.g., countries learn from and emulate their neighbors.
- Clustering: Spatially correlated determinants of welfare benefits and regime type. E.g., states politically dominated by Democrats pay more than those dominated by Republicans, and party dominance clusters regionally; wealthy countries are more likely to be democratic, and there are rich and poor "neighborhoods."

How do we distinguish these possibilities?

Do my outcomes cluster?

- The most popular test for spatial association is Moran's I,

$$I = \frac{N}{S} \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2},$$

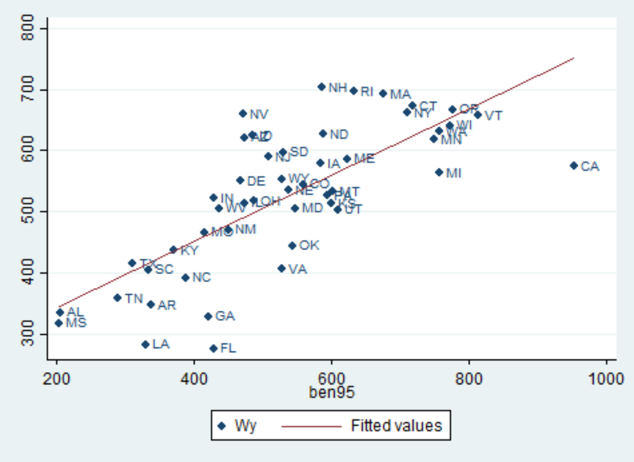
where $S = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}$.

- Or, with OLS residuals

$$I = \frac{N}{S} \frac{\varepsilon' \mathbf{W} \varepsilon}{\varepsilon' \varepsilon}$$

- When **W** is row-standardized, Moran's I is the slope coefficient from the regression of **Wy** on **y**.

Figure: Moran's I for AFDC Benefits



Interdependence vs. Clustering: LM Tests

Now that we know $\mathrm{cov}(\mathbf{y}, \mathbf{W}\mathbf{y}) \neq 0$, how can we identify the source of this covariance? Consider the general model where

$$\begin{aligned} \mathbf{y} &= \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \varepsilon \\ \mathbf{X} &= \phi\mathbf{W}\mathbf{X} + \mathbf{X_0} \\ \varepsilon &= \lambda\mathbf{W}\varepsilon + \mathbf{u} \end{aligned}$$

If we estimate the model

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon,$$

and we assume $\rho = 0$, we can test the restriction that $\lambda = 0$.

$$LM_\lambda = \frac{\left(\hat{\varepsilon}'\mathbf{W}\hat{\varepsilon}/\hat{\sigma}^2_\varepsilon\right)^2}{T},$$

where

$$T = \mathrm{tr}[(\mathbf{W}' + \mathbf{W})\mathbf{W}].$$

Interdependence vs. Clustering: LM Tests

- The problem is that this test has power against the incorrect alternative. If $\rho \neq 0$, under the null hypothesis $\lambda = 0$, $\mathrm{cov}(\hat{\varepsilon}, \mathbf{W}\hat{\varepsilon}) \neq 0$.
- Fortunately, Anselin et al. (1996) have developed a robust LM test for the null hypothesis $\lambda = 0$ that does not make any assumptions about ρ.
- The basic strategy is to remove the portion of the $\mathrm{cov}(\hat{\varepsilon}, \mathbf{W}\hat{\varepsilon})$ that can be attributable to $\mathrm{cov}(\hat{\varepsilon}, \mathbf{Wy})$.

$$LM_{\lambda}^{*} = \frac{\left(\hat{\varepsilon}'\mathbf{W}\hat{\varepsilon}/\hat{\sigma}_{\varepsilon}^{2} - \boldsymbol{\Psi}\hat{\varepsilon}'\mathbf{Wy}/\hat{\sigma}_{\varepsilon}^{2}\right)^{2}}{T\left[1 - \boldsymbol{\Psi}\right]}$$

- A robust LM test for $\rho = 0$ (LM_{ρ}^{*}) can be developed similarly.

Interdependence vs. Clustering: LM Tests

- These tests provide a possible way to distinguish common exposure from diffusion.
- If $\mathrm{cov}(\mathbf{y}, \mathbf{Wy}) \neq 0$ and both LM_{ρ}^{*} and LM_{λ}^{*} fail to reject their respective null hypotheses, one can conclude that the correlation is driven by clustering on **observables**.
- If $\mathrm{cov}(\mathbf{y}, \mathbf{Wy}) \neq 0$, LM_{ρ}^{*} fails to reject and LM_{λ}^{*} rejects, one can conclude that the correlation is driven by clustering on **unobservables**.
- If $\mathrm{cov}(\mathbf{y}, \mathbf{Wy}) \neq 0$, LM_{ρ}^{*} rejects and LM_{λ}^{*} fails to reject, one can conclude that the correlation is driven by outcome **interdependence**.

Maximum Likelihood Estimation

Multivariate change of variables theorem:

$$g(\mathbf{y}) = f\left(r^{-1}(\mathbf{y})\right)|J(\mathbf{y})|$$

The spatial-lag model is:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \varepsilon \Rightarrow \varepsilon = (\mathbf{I} - \rho\mathbf{W})\,\mathbf{y} - \mathbf{X}\beta = \mathbf{A}\mathbf{y} - \mathbf{X}\beta$$

The likelihood for ε is:

$$L(\varepsilon) = \left(\frac{1}{\sigma^2 2\pi}\right)^{N/2} \exp\left(-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right)$$

The inverse function is: $\varepsilon = r^{-1}(\mathbf{y}) = (\mathbf{I} - \rho\mathbf{W})\,\mathbf{y} - \mathbf{X}\beta$

The Jacobian is: $\frac{\partial\varepsilon}{\partial\mathbf{y}} = (\mathbf{I} - \rho\mathbf{W}) = \mathbf{A}$

Thus, the likelihood for $\mathbf{y}$ is

$$L(\mathbf{y}) = |\mathbf{A}| \left(\frac{1}{\sigma^2 2\pi}\right)^{N/2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{A}\mathbf{y} - \mathbf{X}\beta)'\,(\mathbf{A}\mathbf{y} - \mathbf{X}\beta)\right)$$

Calculating Spatial Multipliers

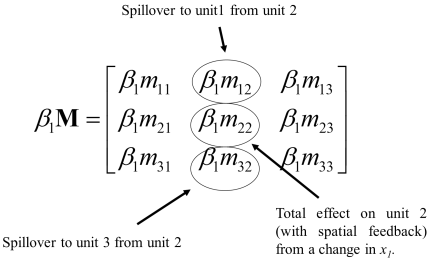
- The spatial lag model is

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X}\beta + \varepsilon$$

- Solving for the reduced-form gives

$$\mathbf{y} = \mathbf{M}(\mathbf{X}\beta + \varepsilon),$$

where $\mathbf{M} = (\mathbf{I} - \rho \mathbf{W})^{-1}$ is the spatial multiplier.



Bootstrapping Confidence Intervals

- Our uncertainty about the spatial multiplier stems from our uncertainty about the estimated parameters $\hat{\beta}$ and $\hat{\rho}$.
- We can generate empirical confidence intervals by sampling from the following bivariate normal distribution.

$$\begin{bmatrix} \beta \\ \rho \end{bmatrix} \sim N\left(\begin{bmatrix} \hat{\beta} \\ \hat{\rho} \end{bmatrix}, \begin{bmatrix} \widehat{\mathrm{var}}(\hat{\beta}) & \widehat{\mathrm{cov}}(\hat{\beta},\hat{\rho}) \\ \widehat{\mathrm{cov}}(\hat{\beta},\hat{\rho}) & \widehat{\mathrm{var}}(\hat{\rho}) \end{bmatrix} \right)$$