10. Comparative Tests among Spatial Regression Models

While the notion of relative likelihood values for different models is somewhat difficult to interpret directly (as mentioned above), such likelihood ratios can in many cases provide powerful *test statistics* for comparing models. In particular, when two models are "nested" in the sense of expression (9.4.1) above, it turns out that the asymptotic distribution of such ratios can be obtained under the (null) hypothesis that the simpler model is the true model. To develop such tests, we begin in Section 10.1 below with a simple one-parameter example where the general ideas to be developed can be given an exact form.

10.1 A One-Parameter Example

Here we revisit the example in Section 8.1 of estimating the mean of a normal random variable, $Y = N(\mu,\sigma^2)$, with known variance, σ^2, given a sample, $y = (y_1,..,y_n)$, of size *n*. The relevant likelihood function is then given by expression (8.1.1) as

$$L(\mu) = L_n(\mu \mid y,\sigma^2) = -n\log\left(\sigma\sqrt{2\pi}\right) - \tfrac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2 \qquad (10.1.1)$$

and the resulting maximum-likelihood estimate of μ, is again seen from expression (8.1.2) to be precisely the sample mean, $\hat{\mu}_n = \overline{y}_n$.

But rather than simply estimating μ, suppose that we now want to test whether $\mu = 0$, or more generally to test the *null hypothesis*, $H_0 : \mu = \mu_0$, for some specified value, μ_0. Then under H_0 the likelihood value in (10.1.1) becomes:

$$L(\mu_0) = L_n(\mu_0 \mid y,\sigma^2) = -n\log\left(\sigma\sqrt{2\pi}\right) - \tfrac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu_0)^2 \qquad (10.1.2)$$

As shown in Figure 10.1 below, it seems reasonable to argue that the likelihood of μ_0 relative to the maximum likelihood at $\hat{\mu}_n$ should provide some indication of the strength

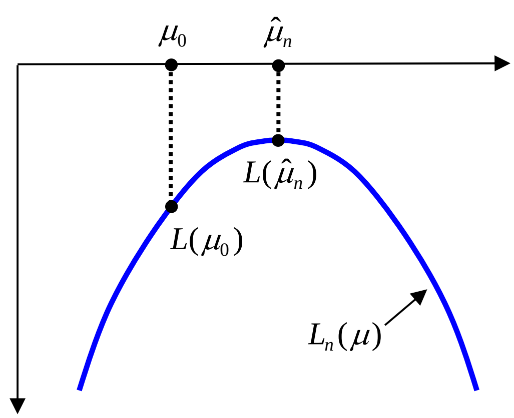


Figure 10.1 Likelihood Comparisons

of evidence in sample y for (or against) hypothesis H_0. In terms of *log* likelihoods, such relations are expressed in terms of the *difference* between $L(\hat{\mu}_n)$ and $L(\mu_0)$. But following standard conventions, we here refer to such log-differences as *likelihood ratios*. Moreover, since $L(\hat{\mu}_n) \geq L(\mu_0)$ by definition, it is natural to focus on the *nonnegative* difference, $L(\hat{\mu}_n) - L(\mu_0)$. If the distribution of $L(\hat{\mu}_n) - L(\mu_0)$ can be determined under H_0, then this statistic can be used to *test* H_0. In particular, if $L(\hat{\mu}_n) - L(\mu_0)$ is "sufficiently large", then this should provide statistical grounds for *rejecting* H_0. With this in mind, observe that by canceling the common terms in the log likelihood expressions, and recalling that $\hat{\mu}_n = \bar{y}_n$, we see that this likelihood ratio can be written as

(10.1.3)
$$\begin{aligned} L(\hat{\mu}_n) - L(\mu_0) &= -\tfrac{1}{2\sigma^2}\left[\sum\nolimits_{i=1}^n (y_i - \hat{\mu}_n)^2 - \sum\nolimits_{i=1}^n (y_i - \mu_0)^2\right] \\ &= -\tfrac{1}{2\sigma^2}\sum\nolimits_{i=1}^n [(y_i - \bar{y}_n)^2 - (y_i - \mu_0)^2] \\ &= -\tfrac{1}{2\sigma^2}\sum\nolimits_{i=1}^n [(y_i^2 - 2y_i\bar{y}_n + \bar{y}_n^2) - (y_i^2 - 2y_i\mu_0 + \mu_0^2)] \\ &= -\tfrac{1}{2\sigma^2}\sum\nolimits_{i=1}^n [-2y_i\bar{y}_n + \bar{y}_n^2 + 2y_i\mu_0 - \mu_0^2] \\ &= -\tfrac{1}{2\sigma^2}\left[-2\bar{y}_n\sum\nolimits_{i=1}^n y_i + n\bar{y}_n^2 + 2\mu_0\sum\nolimits_{i=1}^n y_i - n\mu_0^2\right] \\ &= \tfrac{n}{2\sigma^2}\left[\,\bar{y}_n^2 - 2\mu_0\bar{y}_n + \mu_0^2\,\right] \\ &= \tfrac{n}{2\sigma^2}(\bar{y}_n - \mu_0)^2 \end{aligned}$$

Thus it follows that

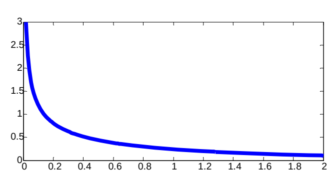
(10.1.4)
$$2[L(\hat{\mu}_n) - L(\mu_0)] = \left(\frac{\bar{y}_n - \mu_0}{\sigma / \sqrt{n}}\right)^2$$

But under the null hypothesis, H_0, the standardized mean in brackets is *standard normal*:

(10.1.5)
$$\frac{\bar{y}_n - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

So the right-hand side of (10.1.4) is distributed as the square of a standard normal variate, which is known to have a *chi square distribution,* χ_1^2, *with one degree of freedom,* i.e.,

(10.1.6)
$$\left(\frac{\bar{y}_n - \mu_0}{\sigma / \sqrt{n}}\right)^2 \sim \chi_1^2$$



where the density of χ_1^2 is plotted on the right. So we may conclude that this *likelihood-ratio statistic* is chi-square distributed (up to a factor of 2) as:

(10.1.7) $$\boxed{2[L(\hat{\mu}_n) - L(\mu_0)] \sim \chi_1^2}$$

[As mentioned in Section 9, this factor of 2 is closely related to the same factor appearing in the penalized likelihood functions developed there.]

Note that we are implicitly comparing two models here, one with a single free parameter (μ) and the other a "nested" special case where μ has been assigned a specific value, μ_0 (typically, $\mu_0 = 0$). But the same likelihood-ratio procedure can be used for much more general comparisons between a "full" model and some special case, denoted as the "restricted" model. Here we simply summarize the main result. Suppose that the *full model* is represented by a log likelihood function, $L(\theta \mid y)$, with parameter vector, $\theta = (\theta_1,..,\theta_K)$, and that the *restricted model* is defined by imposing a set of $m \le K$ restrictions on these parameters that are representable by a vector, $g = (g_j : j = 1,..,m)$, of (smooth) functions as relations of the form,

(10.1.8) $$g_j(\theta) = 0 \,, \quad j = 1,..,m$$

In our simple example above, there is only one relation, namely, $g_1(\mu) = \mu - \mu_0 = 0$. If the maximum-likelihood estimate for *full model* is denoted by $\hat{\theta}$, and if the maximum-likelihood estimate, $\hat{\theta}_g$, for the *restricted model* is taken to be the (unique) solution of the constrained maximization problem,

(10.1.9) $$L(\hat{\theta}_g \mid y) = \max_{\{\theta : g(\theta)=0\}} L(\theta \mid y)$$

then it again follows that the relevant *likelihood-ratio statistic*, $L(\hat{\theta} \mid y) - L(\hat{\theta}_g \mid y)$, is nonnegative. In this more general setting, if it is hypothesized that the restricted model is true (i.e., that the true value of θ satisfies restrictions, g), then under this null hypothesis it can be shown[1] that $L(\hat{\theta} \mid y) - L(\hat{\theta}_g \mid y)$ is now *asymptotically chi square distributed* (up to a factor of 2) with degrees of freedom, m, equal to the number of restrictions defined by g:

(10.1.9) $$\boxed{2[L(\hat{\theta} \mid y) - L(\hat{\theta}_g \mid y)] \sim \chi_m^2}$$

[1] This result, known as *Wilk's Theorem*, is developed, for example, in Section 3.9 of the online *Lecture Notes in Mathematical Statistics* (2003) by R.S. Dudley at MIT (*http://ocw.mit.edu/courses/mathematics/18-466-mathematical-statistics-spring-2003/lecture-notes/*).

This family of *likelihood-ratio tests* provides a general framework for comparing a wide variety of "nested" models. Moreover, as in the one-parameter case of (10.1.7) above, the basic intuition is essentially the same for all such tests. In particular, since the full maximum likelihood, $L(\hat{\theta}_n \,|\, y)$, is almost surely larger than the restricted maximum likelihood, $L(\hat{\theta}_g \,|\, y)$, the only question is whether it is "significantly larger". If so, then it can be argued that the restricted model should be rejected on these grounds. If not, then this suggests that the full model adds little in the way of statistical substance, and thus (by Occam's razor) that the simpler restricted model should be preferred. For example, in the OLS case above, the key question is whether a given parameter, such as β_1, is significantly different from zero (all else being equal). If so, then this indicates that the larger model including variable, x_1, yields a better predictor of y than the same model without x_1.[2] In the following sections, we shall employ this strategy to compare the SE-model and SL-model from a number of perspectives.

10.2 Likelihood-Ratio Tests against OLS

Here we begin by observing that since SEM and SLM are "non-nested" models in the sense that neither is a special case of the other, it is not possible to compare them directly in terms of likelihood-ratio tests. But since OLS is precisely the "$\rho = 0$" case of each model, both SEM and SLM can be compared with OLS in terms of such tests. Thus, by using OLS as a "benchmark" model, we can construct an indirect comparison of SEM and SLM. For example, if the improvement in likelihood of SEM over OLS is much greater than that of SLM over OLS for a given data set, (y, X), then in this sense it can be argued that SEM provides a better model of (y, X) than does SLM.

To operationalize such comparisons, we start with SEM and for a given data set, (y, X), let $(\hat{\beta}_{SEM}, \hat{\sigma}^2_{SEM}, \hat{\rho}_{SEM})$ denote the maximum likelihood estimates obtained using the SEM likelihood function, $L(\beta, \sigma^2, \rho \,|\, y, X)$, in (7.3.4) above [as in expressions (7.3.10) through (7.3.12)]. Then the corresponding *SEM maximum-likelihood value* can be denoted by:

$$\boxed{\hat{L}_{SEM} = L(\hat{\beta}_{SEM}, \hat{\sigma}^2_{SEM}, \hat{\rho}_{SEM} \,|\, y, X)} \qquad (10.2.1)$$

Similarly, if for OLS we let $(\hat{\beta}_{OLS}, \hat{\sigma}^2_{OLS})$ denote the maximum-likelihood estimates in (7.2.6) and (7.2.9) obtained for (y, X) by maximizing (7.2.4), then the corresponding *OLS maximum-likelihood value* can be denoted by

$$\boxed{\hat{L}_{OLS} = L(\hat{\beta}_{OLS}, \hat{\sigma}^2_{OLS} \,|\, y, X)} \qquad (10.2.2)$$

[2] One may ask how this likelihood-ratio test in the OLS case relates to the standard (Wald) tests of significance, such as in expression (8.4.12) above (with $\rho = 0$). Here it can be shown [as for example in Section 13.4 of Davidson and MacKinnon (1993)] that these tests are asymptotically equivalent.

Finally, since the likelihood function in (7.2.4) is clearly the special case of (7.3.4) with $\rho = 0$ [or more precisely, with $g_1(\beta,\sigma^2,\rho) \equiv \rho$ in (10.1.8)], it follows from the general discussion above that under the null hypothesis, $\rho = 0$, it must be true that the *likelihood ratio*, $LR_{SEM/OLS} = 2[\hat{L}_{SEM} - \hat{L}_{OLS}]$, is distributed as chi square with one degree of freedom, i.e., that

(10.2.3) $$\boxed{LR_{SEM/OLS} = 2[\hat{L}_{SEM} - \hat{L}_{OLS}] \sim \chi_1^2}$$

Similarly, if $(\hat{\beta}_{SLM}, \hat{\sigma}^2_{SLM}, \hat{\rho}_{SLM})$ denotes the maximum likelihood estimates obtained using the SLM likelihood function, $L(\beta,\sigma^2,\rho\,|\,y,X)$, in (7.4.2) above [as in expressions (7.4.12) through (7.4.14)], then we may denote the resulting *SLM maximum-likelihood value* by:

(10.2.4) $$\boxed{\hat{L}_{SLM} = L(\hat{\beta}_{SLM}, \hat{\sigma}^2_{SLM}, \hat{\rho}_{SLM}\,|\,y,X)}$$

Then in the same manner as (10.2.3), it follows that under the null hypothesis that $\rho = 0$ for SLM, we also have

(10.2.5) $$\boxed{LR_{SLM/OLS} = 2[\hat{L}_{SLM} - \hat{L}_{OLS}] \sim \chi_1^2}$$

For the Eire case, these two likelihood ratios and associated p-values are reported in Figure 7.7 as

(10.2.6) $$\boxed{LR = LR_{SEM/OLS} = 7.375 \;\; (Pval = .0066)}$$

and

(10.2.7) $$\boxed{LR = LR_{SLM/OLS} = 15.803 \;\; (Pval = .00007)}$$

So for example, if OLS were the correct model, then the chance of obtaining a likelihood ratio, $LR_{SLM/OLS}$, as large as 15.803 would be less than 7 in 100,000. Moreover, while the p-value for $LR_{SEM/OLS}$ is also quite small, it is *relatively less significant* than for SLM. Thus a comparison of these p-values provides at least indirect evidence that SLM is more appropriate than SEM for this Eire data.

But given the indirect nature of this comparison, it is natural to ask whether there are any more direct comparisons. One possibility is developed below, which will be seen to be especially appropriate for the case of row normalized spatial weights matrices.

10.3 The Common-Factor Hypothesis

Here we start by recalling from Section 6.3.2 that if X and β are partitioned as $X = [1_n, X_v]$ and $\beta' = (\beta_0, \beta_v')$, respectively, then an alternative modeling form is provided by the *Spatial Durbin model* (SDM),

$$(10.3.1) \qquad \boxed{Y \;=\; \rho WY \,+\, \beta_0 1_n \,+\, X_v\beta_v \,+ WX_v\alpha \,+\, \varepsilon, \;\; \varepsilon \sim N(0,\sigma^2 I_n)}$$

But this model can be viewed as a special case of the SLM model in the following way. If we group terms in (10.3.1) by letting $X_{SDM} = [1_n, X_v, WX_v]$ and $\beta'_{SDM} = (1'_n, \beta'_v, \alpha')$ so that

$$(10.3.2) \qquad X_{SDM}\beta_{SDM} = [1_n \; X_v \; WX_v]\begin{pmatrix}\beta_0\\ \beta_v\\ \alpha\end{pmatrix} = \beta_0 1_n + X_v\beta_v + WX_v\alpha \;\;,$$

then (10.3.1) can be rewritten as,

$$(10.3.3) \qquad \boxed{Y \;=\; \rho WY \,+\, X_{SDM}\beta_{SDM} \;+\, \varepsilon, \;\; \varepsilon \sim N(0,\sigma^2 I_n)}$$

which is seen to be an instance of SLM in expression (6.2.2).

Moreover, if W is *row normalized*, then SEM can in turn be viewed as a special case of SDM. To see this, observe first that the reduced form of SEM in expression (6.1.9) can be expanded and rewritten as follows:

$$(10.3.4) \qquad Y = X\beta + (I_n - \rho W)^{-1}\varepsilon$$

$$\Rightarrow \; (I_n - \rho W)Y = (I_n - \rho W)X\beta + \varepsilon$$

$$\Rightarrow \; Y - \rho WY = (X - \rho WX)\beta + \varepsilon$$

$$\Rightarrow \; Y = \rho WY + X\beta - \rho WX\beta + \varepsilon$$

So by employing the notation in (10.3.1), we see that

$$(10.3.5) \qquad Y = \rho WY + [\beta_0 1_n + X_v\beta_v] - \rho W[\beta_0 1_n + X_v\beta_v] + \varepsilon$$

$$= \rho WY + \beta_0 1_n + X_v\beta_v - [\rho\beta_0 W1_n + \rho WX_v\beta_v] + \varepsilon$$

Finally, if W is row normalized, then by expression (3.3.30) it follows that $W1_n = 1_n$. So by letting $b_0 = (1-\rho)\beta_0$, and grouping the two unit vector terms, we see finally that the SEM model in (10.3.4) becomes

(10.3.6) $$\boxed{Y = \rho WY + b_0 1_n + X_v \beta_v - WX_v(\rho\beta_v) + \varepsilon}$$

which is precisely SDM in (10.3.1) *under the condition that*

(10.3.7) $$\alpha = -\rho\beta_v$$

This condition is usually formulated as a null hypothesis, designated as the *Common Factor Hypothesis*, and written as

(10.3.8) $$\boxed{H_{CF} : \alpha + \rho\beta_v = 0}$$

Under this hypothesis, it follows that SEM is formally a *restriction* of SDM in the sense of expression (10.1.8), where the relevant vector, *g*, of *restriction functions* is now given by $g(\rho,\beta_0,\beta_v,\alpha,\sigma^2) \equiv \alpha + \rho\beta_v$. The number of restrictions (i.e., *dimension* of *g*) is here simply the number of explanatory variables, k. Given this relationship, one can then employ likelihood-ratio methods to test the appropriateness of SDM versus SEM. To do so for any given any data set, (y, X), we now let $(\hat{\beta}_{SDM}, \hat{\sigma}^2_{SDM}, \hat{\rho}_{SDM})$ denote the maximum likelihood estimates obtained by applying the SLM likelihood function, $L(\beta_{SDM}, \sigma^2, \rho \,|\, y, X)$, in (7.4.2) to the SLM form of SDM in (10.3.3) above. In these terms, the resulting *SDM maximum-likelihood value* is then given by:

(10.3.9) $$\boxed{\hat{L}_{SDM} = L(\hat{\beta}_{SDM}, \hat{\sigma}^2_{SDM}, \hat{\rho}_{SDM} \,|\, y, X)}$$

Finally, if we let $\hat{L}_{SEM} = L(\hat{\beta}_{SEM}, \hat{\sigma}^2_{SEM}, \hat{\rho}_{SEM} \,|\, y, X)$ denote the maximum-likelihood value of the SE-model in (10.3.6) [viewed as an SD-model restricted by (10.3.8)], then under the SEM null hypothesis, we now have

(10.3.10) $$\boxed{LR_{SDM/SEM} = 2[\hat{L}_{SDM} - \hat{L}_{SEM}] \sim \chi^2_k}$$

where again, *k*, is the number of explanatory variables in SEM. The results of this comparative test are part of the SEM output, denoted by ***Com-LR***. For the case of Eire, the result reported in Figure 7.7 is

(10.3.11) $$\boxed{\textit{Com-LR} = 18.427035 \;\; (Pval = 0.000018)}$$

and shows that SDM fits this Blood Group data far better than SEM. This can largely be explained by noting from (10.3.2) and (10.3.3) that the reduced form of the SDM model is given by

(10.3.12) $$Y = B_\rho^{-1}(\beta_0 1_n + X_v\beta_v + WX_v\alpha) + B_\rho^{-1}\varepsilon$$

and thus contains the *Rippled Pale* term, $B_\rho^{-1} X_v \beta_v = (B_\rho^{-1} x)\beta_1$, which was shown to yield a striking fit to this data. So a strong result is not surprising in this case.

Finally, it should be noted that while the above analysis has focused on row-normalized matrices in order to interpret the "SLM version" of SEM as a Spatial Durbin model, this restriction can in principle be relaxed. In particular, when $W1_n \neq 1_n$, it is possible to treat the vector, $W1_n$, as representing the sample values of an additional "explanatory variable" and thus modify (10.3.2) to

(10.3.13) $$X_{SDM}\beta_{SDM} = [1_n \; X_v \; W1_n \; WX_v]\begin{pmatrix}\beta_0\\ \beta_v\\ \alpha_0\\ \alpha\end{pmatrix} = \beta_0 1_n + X_v\beta_v + \alpha_0 W1_n + WX_v\alpha$$

With this addition, SEM can still be viewed formally as an instance of SLM. Moreover, if the additional restriction, $\alpha_0 + \rho\beta_0 = 0$, is added to yield a set of $k+1$ restrictions, then this new likelihood ratio must now be distributed as χ^2_{k+1} under the null hypothesis of SEM. So while the problematic nature of this artificial "explanatory variable" complicates the interpretation of the resulting test, it can still be argued that the presence of the *spatial lag term*, ρWY, suggests that SLM may yield a better fit to the given data than SEM.

10.4 The Combined-Model Approach

A final method of comparing SEM and SLM is provided by the *combined model* (CM) developed in Section 6.3.1 above, which for any given spatial weights matrix, *W*, can be written as [see also expression (6.3.3)]:

(10.4.1) $$\boxed{Y = \rho WY + X\beta + u \;\; , \;\; u = \lambda Wu + \varepsilon \;\; , \;\; \varepsilon \sim N(0,\sigma^2 I_n)}$$

Here is clear that SEM is the special case with $\rho = 0$, and SLM is the special case with $\lambda = 0$. So these two models are seen to lie "between" OLS and the Combined Model, as in Figure 10.2 below:

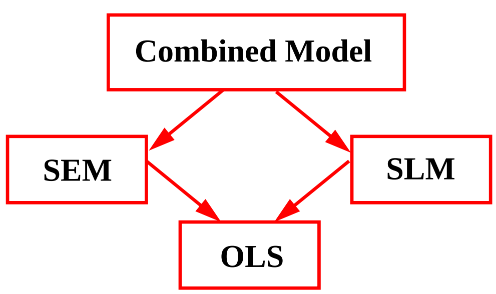


Figure 10.2. Model Relations

In the same way that OLS served as a "lower" benchmark for comparing SEM and SLM, the Combined Model can thus serve as an "upper" benchmark. Here the only issue is how to estimate this more complex model. To do so, we start by observing from (6.3.4) that the reduced form of this model can be written as:

$$\boxed{Y = X_\rho \beta + \varepsilon_{\rho\lambda} \quad , \quad \varepsilon_{\rho\lambda} \sim N(0, \sigma^2 V_{\rho\lambda})} \tag{10.4.2}$$

where

$$X_\rho = (I_n - \rho W)^{-1} X \tag{10.4.3}$$

$$\varepsilon_{\rho\lambda} = (I_n - \rho W)^{-1}(I_n - \lambda W)^{-1} \varepsilon \tag{10.4.4}$$

$$V_{\rho\lambda} = (I_n - \rho W)^{-1}(I_n - \lambda W)^{-1}(I_n - \lambda W')^{-1}(I_n - \rho W')^{-1} \tag{10.4.5}$$

So it should be clear that this model is simply another instance of GLS, where in this case conditioning is on the *pair* of spatial dependence parameters, ρ and λ. So for the parameter vector, $\theta = (\beta, \sigma^2, \rho, \lambda)$, the corresponding likelihood function takes the form:

$$\boxed{L(\theta \mid y) = -\tfrac{n}{2}\log(2\pi) - \tfrac{n}{2}\log(\sigma^2) - \tfrac{1}{2}\log|V_{\rho\lambda}| - \tfrac{1}{2\sigma^2}(y - X_\rho \beta)' V_{\rho\lambda}^{-1}(y - X_\rho \beta)} \tag{10.4.6}$$

and the corresponding conditional maximum-likelihood estimates for β and σ^2 given ρ and λ now take the respective forms:

$$\boxed{\hat{\beta}_{\rho\lambda} = (X_\rho' V_{\rho\lambda}^{-1} X_\rho)^{-1} X_\rho' V_{\rho\lambda}^{-1} y} \tag{10.4.7}$$

$$\boxed{\hat{\sigma}_{\rho\lambda}^2 = \tfrac{1}{n}(y - X_\rho \hat{\beta}_{\rho\lambda})' V_{\rho\lambda}^{-1}(y - X_\rho \hat{\beta}_{\rho\lambda})} \tag{10.4.8}$$

By substituting (10.4.7) and (10.4.8) into (10.4.6), we may then obtain a *concentrated likelihood function* for ρ and λ, denoted by:

$$L_c(\rho, \lambda \mid y) \equiv L(\hat{\beta}_{\rho\lambda}, \hat{\sigma}_{\rho\lambda}^2, \rho, \lambda \mid y) \tag{10.4.9}$$

Finally, by maximizing this two-dimensional function to obtain maximum-likelihood estimates, $\hat{\rho}$ and $\hat{\lambda}$, we can substitute these into (10.4.7) and (10.4.8) to obtain the corresponding maximum-likelihood estimates, $\hat{\beta}_{\hat{\rho}\hat{\lambda}}$ and $\hat{\sigma}_{\hat{\rho}\hat{\lambda}}^2$. This estimation procedure is programed in the MATLAB program, **sac.m**, (**S**patial **A**utocorrelation **C**ombined) written by James Lesage, and can be found in the class directory at:

>> **sys502/Matlab/Lesage_7/spatial/sac_models**

While the parameter estimates, $\hat{\rho}$ and $\hat{\lambda}$, obtained by this procedure often tend to be collinear (in view of their common role in modifying the same weight matrix, *W*), the corresponding *maximum-likelihood value*,

(10.4.10) $$\boxed{\hat{L}_{CM} = L(\hat{\beta}_{\hat{\rho}\hat{\lambda}}, \hat{\sigma}^2_{\hat{\rho}\hat{\lambda}}, \hat{\rho}, \hat{\lambda} \mid y)}$$

continues to be well defined and numerically stable. This value can thus be used to test the relative goodness of fit of the two restricted models, SEM and SLM. In particular, it follows by the same arguments as above that under the *SEM null hypothesis* ($\lambda = 0$) we have

(10.4.11) $$\boxed{\hat{L}_{CM/SEM} = 2[\hat{L}_{CM} - \hat{L}_{SEM}] \sim \chi^2_1}$$

and similarly, that under the *SLM null hypothesis* ($\rho = 0$) we have

(10.4.11) $$\boxed{\hat{L}_{CM/SLM} = 2[\hat{L}_{CM} - \hat{L}_{SLM}] \sim \chi^2_1}$$

The results of these respective tests for the Eire case are as follows:

(10.4.12) $$\boxed{LR_{CM/SEM} = 10.92 \quad (Pval = .0009)}$$

(10.4.13) $$\boxed{LR_{CM/SLM} = 2.49 \quad (Pval = .1145)}$$

Thus the Combined Model is seen to yield a significantly better fit than SEM, but *not* SLM. So relative to this CM benchmark, it can again be concluded that SLM yields a better fit to the Eire data than does SEM.