ΚΕΦΑΛΑΙΟ 5

ΧΩΡΙΚΑ ΟΙΚΟΝΟΜΕΤΡΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ

5.1 Εισαγωγή

Η εκτίμηση υποδειγμάτων γραμμικής παλινδρόμησης με χωρικά δεδομένα, δηλαδή με δεδομένα που προέρχονται από διαφορετικές γεωγραφικές περιοχές, πολλές φορές υποκρύπτει τον κίνδυνο εμφάνισης του προβλήματος της χωρικής αυτοσυσχέτισης στα σφάλματα. Πάρα πολλοί ερευνητές αγνοούν τον κίνδυνο αυτό και θεωρούν ότι λόγω των διαστρωματικών δεδομένων η ετεροσκεδαστικότητα των σφαλμάτων είναι ένα πιθανό πρόβλημα που μπορεί να προκύψει και ότι η αυτοσυσχέτιση εμφανίζεται μόνο όταν το δείγμα αποτελείται από χρονοσειρές. Η χωρική αυτοσυσχέτιση (*spatial autocorrelation*) είναι αποτέλεσμα της χωρικής εξάρτησης (*spatial dependence*), η οποία μαζί με τη χωρική ετερογένεια (*spatial heterogeneity*) αποτελούν τις χωρικές επιδράσεις (*spatial effects*) που οφείλονται στο είδος των δεδομένων. Η ύπαρξη χωρικής αυτοσυσχέτισης, δηλαδή συσχέτισης των παρατηρήσεων σύμφωνα με τις γεωγραφικές τους θέσεις, καθιστά τα αποτελέσματα από την εκτίμηση ενός οικονομετρικού υποδείγματος αναξιόπιστα. Κατά συνέπεια, ο ερευνητής θα πρέπει να είναι σε θέση να εντοπίσει και να αντιμετωπίσει το πρόβλημα αυτό, εκτιμώντας χωρικά οικονομετρικά υποδείγματα που περιλαμβάνουν χωρικές επιδράσεις.

Στο κεφάλαιο αυτό παρουσιάζονται οι πιο συνηθισμένες αιτίες που προκαλούν την εμφάνιση του προβλήματος της χωρικής αυτοσυσχέτισης στα σφάλματα ενός γραμμικού οικονομετρικού υποδείγματος παλινδρόμησης και οι αντίστοιχες συνέπειες στη στατιστική συμπερασματολογία της μεθόδου των ελαχίστων τετραγώνων. Στη συνέχεια, εξετάζονται τα δύο βασικά χωρικά οικονομετρικά υποδείγματα που εκτιμώνται όταν εντοπίζεται χωρική αυτοσυσχέτιση, δηλαδή το υπόδειγμα χωρικού σφάλματος και το υπόδειγμα χωρικής υστέρησης. Τέλος, αναλύονται οι πιο διαδομένοι έλεγχοι που χρησιμοποιούνται στην ανίχνευση του προβλήματος και στην επιλογή του κατάλληλου χωρικού υποδείγματος.

5.2 Οι συνέπειες της χωρικής αυτοσυσχέτισης στα σφάλματα του γραμμικού υποδείγματος παλινδρόμησης

Στην οικονομετρική ανάλυση η μέθοδος της ανάλυσης παλινδρόμησης έχει σκοπό τη διερεύνηση της σχέσης ανάμεσα σε μια εξαρτημένη μεταβλητή y και ένα σύνολο ανεξάρτητων μεταβλητών $x_1, x_2, \ldots, x_k$. Συνήθως, θεωρείται ότι η σχέση αυτή είναι γραμμική, οπότε ένα υπόδειγμα που θα μπορούσε να χρησιμοποιηθεί για αυτό το σκοπό είναι το ακόλουθο:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i \quad \text{για} \quad i = 1, 2, \ldots, n$$

το οποίο με τη βοήθεια της άλγεβρας μητρών γράφεται ως εξής:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{5.1}$$

όπου $\mathbf{y}$ είναι το ($n \times 1$) διάνυσμα των παρατηρήσεων της εξαρτημένης μεταβλητής, $\mathbf{X}$ η ($n \times (k+1)$) μήτρα των παρατηρήσεων των k ανεξάρτητων μεταβλητών έχοντας στην πρώτη στήλη μονάδες για την εκτίμηση του σταθερού όρου, $\boldsymbol{\beta}$ το ($(k+1) \times 1)$) διάνυσμα των συντελεστών του υποδείγματος και $\boldsymbol{\varepsilon}$ το ($n \times 1$) διάνυσμα των τυχαίων σφαλμάτων που περιλαμβάνουν όλους τους παράγοντες που επηρεάζουν την εξαρτημένη μεταβλητή και δεν έχουν συμπεριληφθεί στις ανεξάρτητες μεταβλητές του υποδείγματος. Το υπόδειγμα της σχέσης (5.1) δεν λαμβάνει υπ' όψιν του την παρουσία χωρικών επιδράσεων και για το λόγο αυτό θα ονομάζεται απλό οικονομετρικό υπόδειγμα για να διακρίνεται από τα χωρικά υποδείγματα.

Επιπρόσθετα, για την εκτίμηση του παραπάνω υποδείγματος με τη μέθοδο των ελαχίστων τετραγώνων (OLS) χρησιμοποιούνται οι ακόλουθες υποθέσεις:

1) $\mathrm{E}[\boldsymbol{\varepsilon}] = \mathbf{0}$, δηλαδή τα σφάλματα έχουν μέσο όρο μηδέν.
2) $\mathrm{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \sigma^2\mathbf{I}$, δηλαδή τα σφάλματα έχουν σταθερή διακύμανση και είναι μεταξύ τους ανεξάρτητα, όπου με $\mathbf{I}$ συμβολίζεται η μοναδιαία μήτρα.
3) $\mathrm{E}[\mathbf{X}'\boldsymbol{\varepsilon}] = \mathbf{0}$, δηλαδή τα σφάλματα δεν συσχετίζονται με τις ανεξάρτητες μεταβλητές.
4) $r[\mathbf{X}] = k \;\Rightarrow |\mathbf{X}'\mathbf{X}| \neq 0$, δηλαδή η μήτρα $\mathbf{X}$ είναι πλήρους τάξεως, που σημαίνει ότι οι ανεξάρτητες μεταβλητές δεν συσχετίζονται γραμμικά μεταξύ τους.
5) $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, δηλαδή τα σφάλματα ακολουθούν την πολυμεταβλητή κανονική κατανομή με μέσο όρο μηδέν και σταθερή διακύμανση.

Όταν δεν ισχύει η υπόθεση της σταθερής διακύμανσης των σφαλμάτων εμφανίζεται το πρόβλημα της ετεροσκεδαστικότητας, όταν παραβιάζεται η υπόθεση της ανεξαρτησίας των σφαλμάτων εμφανίζεται το πρόβλημα της αυτοσυσχέτισης και όταν οι ανεξάρτητες μεταβλητές συσχετίζονται γραμμικά εμφανίζεται το πρόβλημα της πολυσυγραμμικότητας. Εφόσον τηρούνται όλες οι προηγούμενες υποθέσεις, οι εκτιμητές των συντελεστών του υποδείγματος (5.1) που προκύπτουν από την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων είναι οι $\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{\mathbf{-1}} \mathbf{X}'\mathbf{y}$ που σύμφωνα με το θεώρημα Gauss-Markov ονομάζονται BLUE (Best Linear Unbiased Estimators) διότι είναι γραμμικοί ως προς τις τιμές της εξαρτημένης μεταβλητής, αμερόληπτοι και άριστοι, δηλαδή έχουν τη μικρότερη διακύμανση από όλους τους γραμμικούς και αμερόληπτους εκτιμητές των συντελεστών του υποδείγματος. Η διακύμανση των εκτιμητών σε αυτήν την περίπτωση ορίζεται ως $\mathrm{Var}\left[\hat{\boldsymbol{\beta}}\right] = \sigma^2 \left(\mathbf{X}'\mathbf{X}\right)^{\mathbf{-1}}$.

Όταν οι τιμές των σφαλμάτων του υποδείγματος (5.1) δεν είναι ανεξάρτητες αλλά παρουσιάζουν κάποια μορφή χωρικής αυτοσυσχέτισης, τότε αποδεικνύεται ότι η μέθοδος των ελαχίστων τετραγώνων θα οδηγήσει σε εκτιμητές μη αποδοτικούς και ανάλογα με την αιτία δημιουργίας του προβλήματος μπορεί να είναι αμερόληπτοι ή μεροληπτικοί. Ειδικότερα, το πρόβλημα της χωρικής αυτοσυσχέτισης στα σφάλματα ενός απλού οικονομετρικού υποδείγματος μπορεί να εμφανίζεται ως ενόχληση ή με ουσιαστική μορφή. Η πρώτη περίπτωση που έχει ηπιότερες συνέπειες, όπως συμβαίνει και στην χρονική αυτοσυσχέτιση, αντιμετωπίζεται εκτιμώντας το υπόδειγμα χωρικού σφάλματος. Αντίθετα, η δεύτερη περίπτωση έχει σοβαρότερες συνέπειες καθώς δηλώνει την παρουσία χωρικής εξάρτησης στην εξαρτημένη μεταβλητή του υποδείγματος που δεν μπορεί να μη ληφθεί υπ’ όψιν στην εξειδίκευσή του διότι διαφορετικά θα προκύψουν παραπλανητικά αποτελέσματα. Όταν η χωρική αυτοσυσχέτιση είναι ουσιαστική η αντιμετώπισή της γίνεται με την εκτίμηση του υποδείγματος χωρικής υστέρησης. Στη συνέχεια αυτής της ενότητας θα παρουσιαστούν αναλυτικά οι συνέπειες της κάθε αιτίας χωρικής αυτοσυσχέτισης για το γραμμικό υπόδειγμα παλινδρόμησης.

5.2.1 Οι συνέπειες από την εμφάνιση χωρικής αυτοσυσχετισης ως ενόχληση

Για να εξεταστούν οι συνέπειες στο γραμμικό υπόδειγμα παλινδρόμησης που προκαλούνται από την παρουσία χωρικής αυτοσυσχέτισης στα σφάλματα ως ενόχληση θα θεωρηθεί αρχικά ότι τα σφάλματα σχηματίζονται σύμφωνα με μια στάσιμη χωρική αυτοπαλίνδρομη διαδικασία πρώτης τάξης (*Spatial Autoregressive Process*–SAR(1)), δηλαδή ως εξής:

$$\boldsymbol{\varepsilon} = \lambda \mathbf{W}\boldsymbol{\varepsilon} + \mathbf{u} \tag{5.2}$$

όπου $\mathbf{u} \sim N(0, \sigma^2 \mathbf{I})$, λ είναι ο χωρικά αυτοπαλίνδρομος συντελεστής για τον οποίο γίνεται η υπόθεση ότι $1/\omega_{\min} < \lambda < 1/\omega_{\max}$ και $\mathbf{W}$ είναι η ($n \times n$) μήτρα χωρικών σταθμίσεων με $\omega_{\min}$ και $\omega_{\max}$ τη μικρότερη και μεγαλύτερη ιδιοτιμή της. Η σχέση (5.2) ισοδύναμα γράφεται ως:

$$\boldsymbol{\varepsilon} = \left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathbf{u} \tag{5.3}$$

οπότε το απλό γραμμικό υπόδειγμα γίνεται:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \Rightarrow \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathbf{u}$$

Δηλαδή, πρόκειται για ένα απλό οικονομετρικό υπόδειγμα που όμως τα σφάλματά του δεν είναι ασυσχέτιστα. Ας υποτεθεί ότι το υπόδειγμα αυτό εκτιμάται με τη μέθοδο των ελαχίστων τετραγώνων (OLS) από κάποιον ερευνητή που αγνοεί τον τρόπο σχηματισμού των σφαλμάτων του.

Η μέση τιμή των σφαλμάτων του υποδείγματος σύμφωνα με τη σχέση (5.3) θα είναι:

$$\mathrm{E}\left[\boldsymbol{\varepsilon}\right] = \mathrm{E}\left[\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathbf{u}\right] = \left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathrm{E}\left[\mathbf{u}\right] = 0$$

Κατά συνέπεια, τα σφάλματα από την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων στο υπόδειγμα (5.1) θα εξακολουθούν να έχουν μέση τιμή μηδέν.

Η μήτρα διακυμάνσεων – συνδιακυμάνσεων των σφαλμάτων θα είναι:

$$\mathrm{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] = \mathrm{E}\left[\left(\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathbf{u}\right)\left(\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathbf{u}\right)'\right] = \mathrm{E}\left[\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathbf{u}\mathbf{u}'\left(\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1}\right)'\right]$$

$$= \left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \mathrm{E}\left[\mathbf{u}\mathbf{u}'\right]\left(\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1}\right)' = \sigma^2 \left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} \left(\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1}\right)' \neq \sigma^2 \mathbf{I}$$

Δηλαδή, παραβιάζεται η δεύτερη υπόθεση του απλού γραμμικού υποδείγματος. Μόνο όταν $\lambda = 0$ θα ικανοποιείται η υπόθεση $\mathrm{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] = \sigma^2 \mathbf{I}$. Πιο συγκεκριμένα, επειδή ισχύει ότι:

$$\left(\mathbf{I} - \lambda \mathbf{W}\right)^{-1} = \mathbf{I} + \lambda \mathbf{W} + \lambda^2 \mathbf{W}^2 + \lambda^3 \mathbf{W}^3 + \ldots$$

η μήτρα διακυμάνσεων – συνδιακυμάνσεων των τιμών του τυχαίου σφάλματος γράφεται αναλυτικά ως:

$$\mathrm{E}\left[\boldsymbol{\varepsilon\varepsilon}'\right]=\sigma^2\left[\mathbf{I}+\lambda\left(\mathbf{W}+\mathbf{W}'\right)+\lambda^2\left(\mathbf{W}^2+\mathbf{WW}'+\mathbf{W}'^2\right)+\ldots\ldots\right]$$

Επομένως, διαπιστώνεται ότι εμφανίζεται πρόβλημα αυτοσυσχέτισης στα σφάλματα τα οποία συσχετίζονται σύμφωνα με τη γεωγραφική τους θέση, όπως αυτή καθορίζεται από τη μήτρα $\mathbf{W}$, κατά τέτοιον τρόπο ώστε κάθε περιοχή να συσχετίζεται με όλες τις άλλες περιοχές αλλά οι περιοχές που βρίσκονται χωρικά πιο κοντά μεταξύ τους να παρουσιάζουν πιο έντονη συσχέτιση από τις πιο απομακρυσμένες. Ταυτόχρονα, παρατηρείται ότι τα στοιχεία στην κύρια διαγώνιο της μήτρας διακυμάνσεων – συνδιακυμάνσεων δεν είναι σταθερά που σημαίνει ότι εμφανίζεται και το πρόβλημα της ετεροσκεδαστικότητας κάτι που περιπλέκει περισσότερο την ανάλυση. Αυτό έρχεται σε άμεση αντίθεση με την περίπτωση της παλινδρόμησης όπου το δείγμα αποτελείται από χρονοσειρές στην οποία όταν τα σφάλματα σχηματίζονται σύμφωνα με μια αυτοπαλίνδρομη διαδικασία εμφανίζεται μόνο το πρόβλημα της αυτοσυσχέτισης και όχι και της ετεροσκεδαστικότητας.

Για τους εκτιμητές των συντελεστών του υποδείγματος προκύπτει ότι ισχύει:

$$\begin{aligned}\mathrm{E}\left[\hat{\boldsymbol{\beta}}\right]&=\mathrm{E}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}\right]=\mathrm{E}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\left(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}\right)\right]\\&=\mathrm{E}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}+\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right]\\&=\mathrm{E}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}+\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\left(\mathbf{I}-\lambda\mathbf{W}\right)^{-1}\mathbf{u}\right]\\&=\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{X}\mathrm{E}\left[\boldsymbol{\beta}\right]+\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\left(\mathbf{I}-\lambda\mathbf{W}\right)^{-1}\mathrm{E}\left[\mathbf{u}\right]=\boldsymbol{\beta}+\mathbf{0}=\boldsymbol{\beta}\end{aligned}$$

και έτσι, οι εκτιμητές από την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων παραμένουν αμερόληπτοι.

Η μήτρα διακυμάνσεων – συνδιακυμάνσεων των εκτιμητών του υποδείγματος θα είναι:

$$\begin{aligned}\mathrm{Var}\left[\hat{\boldsymbol{\beta}}\right]&=\mathrm{E}\left[\left(\hat{\boldsymbol{\beta}}-\mathrm{E}\left[\hat{\boldsymbol{\beta}}\right]\right)\left(\hat{\boldsymbol{\beta}}-\mathrm{E}\left[\hat{\boldsymbol{\beta}}\right]\right)'\right]\\&=\mathrm{E}\left[\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\right]\\&=\mathrm{E}\left[\left(\boldsymbol{\beta}+\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\varepsilon}-\boldsymbol{\beta}\right)\left(\boldsymbol{\beta}+\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\varepsilon}-\boldsymbol{\beta}\right)'\right]\end{aligned}$$

$$=\mathrm{E}\left[\left((\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\boldsymbol{\varepsilon}\right)\left((\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\boldsymbol{\varepsilon}\right)'\right]$$

$$=\mathrm{E}\left[\left((\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\boldsymbol{\varepsilon}\right)\left(\boldsymbol{\varepsilon}'\mathbf{X}\left((\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\right)'\right)\right]$$

$$=\mathrm{E}\left[(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\right]$$

$$=(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\mathrm{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']\mathbf{X}(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}$$

$$=(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\sigma^2(\mathbf{I}-\lambda\mathbf{W})^{-1}\left((\mathbf{I}-\lambda\mathbf{W})^{-1}\right)'\mathbf{X}(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}$$

Διότι, επειδή η μήτρα $(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}$ είναι συμμετρική θα ισχύει $\left((\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\right)'=(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}$. Κατά συνέπεια, διαπιστώνεται ότι $\mathrm{Var}\left[\hat{\boldsymbol{\beta}}\right]\neq\sigma^2(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}$ που σημαίνει ότι οι εκτιμητές που προκύπτουν από την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων παύουν να είναι αποδοτικοί, καθώς η μήτρα διακυμάνσεων – συνδιακυμάνσεων των εκτιμητών λαμβάνει μια αρκετά σύνθετη μορφή που καθιστά μη αξιόπιστη τη στατιστική συμπερασματολογία και την εφαρμογή των στατιστικών ελέγχων.

Μια άλλη στοχαστική διαδικασία που μπορεί να θεωρηθεί ότι ακολουθούν τα σφάλματα του υποδείγματος (5.1) είναι η χωρική διαδικασία κινητού μέσου πρώτης τάξης (*Spatial Moving Average Process*–SMA(1)), η οποία ορίζεται ως:

$$\boldsymbol{\varepsilon}=\lambda\mathbf{W}\mathbf{u}+\mathbf{u}$$

ή ισοδύναμα ως:

$$\boldsymbol{\varepsilon}=(\mathbf{I}+\lambda\mathbf{W})\mathbf{u} \tag{5.4}$$

όπου για το λ υποτίθεται ότι $1/\omega_{\min}<-\lambda<1/\omega_{\max}$. Το απλό γραμμικό υπόδειγμα σε αυτή την περίπτωση γίνεται:

$$\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon} \quad\Rightarrow\quad \mathbf{y}=\mathbf{X}\boldsymbol{\beta}+(\mathbf{I}+\lambda\mathbf{W})\mathbf{u}$$

Εύκολα μπορεί να αποδειχθεί ότι και στην περίπτωση του χωρικού κινητού μέσου τα σφάλματα του απλού οικονομετρικού υποδείγματος θα έχουν μέση τιμή μηδέν αφού σύμφωνα με τη σχέση (5.4) θα ισχύει:

$$\mathrm{E}[\boldsymbol{\varepsilon}]=\mathrm{E}\left[(\mathbf{I}+\lambda\mathbf{W})\mathbf{u}\right]=(\mathbf{I}+\lambda\mathbf{W})\mathrm{E}[\mathbf{u}]=0$$

Η μήτρα διακυμάνσεων – συνδιακυμάνσεων των σφαλμάτων θα είναι:

$$\mathrm{E}[\boldsymbol{\varepsilon\varepsilon}'] = \mathrm{E}\left[\left((\mathbf{I}+\lambda\mathbf{W})\mathbf{u}\right)\left((\mathbf{I}+\lambda\mathbf{W})\mathbf{u}\right)'\right] = \mathrm{E}\left[(\mathbf{I}+\lambda\mathbf{W})\mathbf{uu}'(\mathbf{I}+\lambda\mathbf{W}')\right]$$
$$= \mathrm{E}[\mathbf{uu}'](\mathbf{I}+\lambda\mathbf{W})(\mathbf{I}+\lambda\mathbf{W}') = \sigma^2\left(\mathbf{I}+\lambda(\mathbf{W}+\mathbf{W}')+\lambda^2\mathbf{WW}'\right) \neq \sigma^2\mathbf{I}$$

και κατά συνέπεια, όπως και η χωρική αυτοπαλίνδρομη διαδικασία, που εξετάστηκε προηγουμένως, έτσι και η χωρική διαδικασία κινητού μέσου θα επιφέρει το πρόβλημα της αυτοσυσχέτισης και της ετεροσκεδαστικότητας στα σφάλματα του υποδείγματος (5.1). Η ουσιαστική διαφορά ανάμεσα στις δύο διαδικασιές είναι ότι η χωρική αυτοσυσχέτιση που προκαλείται από τη διαδικασία κινητού μέσου είναι πιο ασθενής, αφού η αλληλεπίδραση μεταξύ των περιοχών εμφανίζεται μόνο μεταξύ των πρώτων και δεύτερων γειτόνων σε αντίθεση με την αυτοπαλίνδρομη διαδικασία που όλες οι γεωγραφικές περιοχές αλληλοεξαρτώνται. Ανάλογα αποδεικνύεται ότι οι εκτιμητές του υποδείγματος θα είναι και πάλι αμερόληπτοι αφού προκύπτει:

$$\mathrm{E}\left[\hat{\boldsymbol{\beta}}\right] = \mathrm{E}\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}\right] = \mathrm{E}\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon})\right]$$
$$= \mathrm{E}\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right]$$
$$= \mathrm{E}\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{I}+\lambda\mathbf{W})\mathbf{u}\right]$$
$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\mathrm{E}[\boldsymbol{\beta}]+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{I}+\lambda\mathbf{W})\mathrm{E}[\mathbf{u}] = \boldsymbol{\beta}+\mathbf{0} = \boldsymbol{\beta}$$

Ωστόσο, αν και αμερόληπτοι οι εκτιμητές δεν θα είναι αποδοτικοί καθώς η διακύμανσή τους θα είναι ίση με:

$$\mathrm{Var}\left[\hat{\boldsymbol{\beta}}\right] = \mathrm{E}\left[\left(\hat{\boldsymbol{\beta}}-\mathrm{E}\left[\hat{\boldsymbol{\beta}}\right]\right)\left(\hat{\boldsymbol{\beta}}-\mathrm{E}\left[\hat{\boldsymbol{\beta}}\right]\right)'\right]$$
$$= \mathrm{E}\left[\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\right]$$
$$= \mathrm{E}\left[\left(\boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}-\boldsymbol{\beta}\right)\left(\boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}-\boldsymbol{\beta}\right)'\right]$$
$$= \mathrm{E}\left[\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right)\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right)'\right]$$
$$= \mathrm{E}\left[\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right)\left(\boldsymbol{\varepsilon}'\mathbf{X}\left((\mathbf{X}'\mathbf{X})^{-1}\right)'\right)\right]$$
$$= \mathrm{E}\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\right]$$
$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{E}[\boldsymbol{\varepsilon\varepsilon}']\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$
$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2(\mathbf{I}+\lambda\mathbf{W})(\mathbf{I}+\lambda\mathbf{W}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \neq \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

Όπως αναφέρουν οι Anselin και Bera (1998) κάθε διαδικασία που εμπλέκει χωρική εξάρτηση στα σφάλματα θα οδηγεί στην εμφάνιση των προβλημάτων της αυτοσυσχέτισης και της ετεροσκεδαστικότητας. Μοναδική εξαίρεση που δεν θα προκληθεί ετεροσκεδαστικότητα αποτελεί η περίπτωση στην οποία τα στοιχεία της μήτρας **W** έχουν οριστεί με τέτοιο τρόπο ώστε όλες οι παρατηρήσεις να έχουν τον ίδιο αριθμό γειτόνων και ίδιες σταθμίσεις. Επομένως, οι εκτιμητές που προκύπτουν από την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων θα προσδιορίζονται σωστά αφού παραμένουν αμερόληπτοι αλλά η διακύμανσή τους θα προσδιορίζεται λανθασμένα.

5.2.2 Οι συνέπειες από την εμφάνιση ουσιαστικής χωρικής αυτοσυσχετισης

Έστω ότι η πραγματική σχέση ανάμεσα στην εξαρτημένη μεταβλητή και τις ανεξάρτητες μεταβλητές είναι της μορφής:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{5.5}$$

όπου **Wy** είναι η χωρική υστέρηση της εξαρτημένης μεταβλητής και ρ ο συντελεστής χωρικής υστέρησης για τον οποίο ισχύει ότι $1/\omega_{\min} < \rho < 1/\omega_{\max}$. Στη σχέση του υποδείγματος (5.5) η εξαρτημένη μεταβλητή **y** σε κάθε περιοχή επηρεάζεται από τις τιμές της στις γειτονικές της περιοχές. Εάν υποτεθεί ότι ο ερευνητής αγνοεί την παρουσία της χωρικής υστέρησης στην εξαρτημένη μεταβλητή και εκτιμήσει με τη μέθοδο των ελαχίστων τετραγώνων το απλό οικονομετρικό υπόδειγμα που ορίστηκε στη σχέση (5.1), τότε η αναμενόμενη τιμή του διανύσματος των εκτιμητών είναι:

$$\begin{aligned}
\mathrm{E}\left[\hat{\boldsymbol{\beta}}\right] &= \mathrm{E}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}\right] = \mathrm{E}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\left(\rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\right)\right] \\
&= \mathrm{E}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\rho\mathbf{W}\mathbf{y} + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right] \\
&= \mathrm{E}\left[\boldsymbol{\beta}\right] + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}\mathrm{E}\left[\rho\right] + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathrm{E}\left[\boldsymbol{\varepsilon}\right] \\
&= \boldsymbol{\beta} + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}\rho = \boldsymbol{\beta} + \mathbf{b}_{\mathbf{Wy},\mathbf{X}}\rho
\end{aligned}$$

όπου με $\mathbf{b}_{\mathbf{Wy},\mathbf{X}}$ συμβολίζεται το διάνυσμα των εκτιμητών από την εκτίμηση με τη μέθοδο των ελαχίστων τετραγώνων του γραμμικού υποδείγματος που έχει ως εξαρτημένη μεταβλητή τη χωρική υστέρηση της εξαρτημένης μεταβλητής και ανεξάρτητες μεταβλητές τις υπόλοιπες ανεξάρτητες μεταβλητές του υποδείγματος (5.5). Επομένως, είναι φανερό ότι οι εκτιμητές που προκύπτουν είναι μη συνεπείς και μεροληπτικοί κάτι που συνεπάγεται ότι και η διακύμανσή τους δεν θα προσδιοριστεί σωστά. Ειδικότερα, εάν θεωρηθεί ότι η εξαρτημένη και οι ανεξάρτητες μεταβλητές αποτελούνται από θετικές τιμές κάτι που συνήθως ισχύει για τα οικονομικά δεδομένα και ο συντελεστής ρ είναι θετικός οι τιμές των εκτιμητών που προκύπτουν θα είναι μεγαλύτερες από τις πραγματικές. Επιπρόσθετα, η παράλειψη της χωρικής υστέρησης της εξαρτημένης μεταβλητής από τις ερμηνευτικές μεταβλητές του υποδείγματος θα οδηγήσει στην εμφάνιση του φαινομένου της χωρικής αυτοσυσχέτισης στα σφάλματα του απλού οικονομετρικού υποδείγματος για τον ίδιο λόγο που εμφανίζεται η χρονική αυτοσυσχέτιση στα σφάλματα ενός οικονομετρικού υποδείγματος που εκτιμάται με χρονοσειρές όταν παραλείπεται η χρονική υστέρηση της εξαρτημένης μεταβλητής εφόσον αυτή συνεισφέρει στην ερμηνεία του υποδείγματος. Αυτό φαίνεται εύκολα καθώς η παράλειψη της χωρικής υστέρησης συνεπάγεται την ενσωμάτωσή της στα σφάλματα του υποδείγματος που θα ορίζονται ως $\mathbf{v} = \rho\mathbf{Wy} + \boldsymbol{\varepsilon}$, οπότε η μήτρα διακυμάνσεων – συνδιακυμάνσεων των σφαλμάτων θα είναι:

$$\begin{aligned} \mathrm{E}\left[\mathbf{vv}'\right] &= \mathrm{E}\left[\left(\rho\mathbf{Wy}+\boldsymbol{\varepsilon}\right)\left(\rho\mathbf{Wy}+\boldsymbol{\varepsilon}\right)'\right] = \mathrm{E}\left[\left(\rho\mathbf{Wy}+\boldsymbol{\varepsilon}\right)\Big(\rho\left(\mathbf{Wy}\right)'+\boldsymbol{\varepsilon}'\Big)\right] \\ &= \mathrm{E}\left[\rho^2\left(\mathbf{Wy}\right)\left(\mathbf{Wy}\right)' + \rho\left(\mathbf{Wy}\right)\boldsymbol{\varepsilon}' + \rho\boldsymbol{\varepsilon}\left(\mathbf{Wy}\right)' + \boldsymbol{\varepsilon\varepsilon}'\right] \neq \sigma^2\mathbf{I} \end{aligned}$$

στην οποία οι τιμές των στοιχείων εκτός της κυρίας διαγωνίου δεν είναι μηδέν. Μόνο όταν $\rho = 0$ θα ικανοποιείται η υπόθεση $\mathrm{E}\left[\boldsymbol{\varepsilon\varepsilon}'\right] = \sigma^2\mathbf{I}$.

Μία περίπτωση που παρουσιάζει ιδιαίτερο ενδιαφέρον είναι όταν ο ερευνητής γνωρίζει για την εμφάνιση της χωρικής υστέρησης στην εξαρτημένη μεταβλητή και επιδιώκει να εκτιμήσει το υπόδειγμα (5.5) με τη βοήθεια της μεθόδου των ελαχίστων τετραγώνων. Όταν συμβεί αυτό, μπορεί να διαπιστωθεί ότι θα παραβιαστεί η τρίτη υπόθεση της μεθόδου των ελαχίστων τετραγώνων $\mathrm{E}\left[\mathbf{X}'\boldsymbol{\varepsilon}\right] = \mathbf{0}$, για τη χωρική υστέρηση $\mathbf{Wy}$ που βρίσκεται ως ανεξάρτητη μεταβλητή στη δεξιά πλευρά του υποδείγματος οδηγώντας σε πρόβλημα ενδογένειας. Διαισθητικά αυτό φαίνεται ως εξής:

$$\mathbf{y} = \rho\mathbf{Wy} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \;\Rightarrow\; \mathbf{y} = \left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}\mathbf{X}\boldsymbol{\beta} + \left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}\boldsymbol{\varepsilon}$$

και πολλαπλασιάζοντας και τα δύο μέλη της εξίσωσης με τη μήτρα χωρικών σταθμίσεων $\mathbf{W}$ προκύπτει:

$$\mathbf{Wy} = \mathbf{W}\left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}\boldsymbol{\varepsilon}$$

από όπου φανερώνεται ότι η μήτρα $\mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}$ δεν έχει όλα τα διαγώνια στοιχεία ίσα με μηδέν οπότε κάθε στοιχείο της **Wy** θα εξαρτάται από την ίδια γραμμή στο διάνυσμα **ε**.

Επιπρόσθετα, η συσχέτιση μεταξύ της χωρικής υστέρησης **Wy** και του σφάλματος **ε** μπορεί να εκφραστεί ως εξής:

$$
\begin{aligned}
\mathrm{E}\left[(\mathbf{Wy})\boldsymbol{\varepsilon}'\right] &= \mathrm{E}\left[\mathbf{W}\left((\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}+(\mathbf{I}-\rho\mathbf{W})^{-1}\boldsymbol{\varepsilon}\right)\boldsymbol{\varepsilon}'\right] \\
&= \mathrm{E}\left[\left(\mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}+\mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\boldsymbol{\varepsilon}\right)\boldsymbol{\varepsilon}'\right] \\
&= \mathrm{E}\left[\mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}\boldsymbol{\varepsilon}'+\mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] \\
&= \mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}\mathrm{E}\left[\boldsymbol{\varepsilon}'\right]+\mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\mathrm{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] \\
&= \mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\mathrm{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] \\
&= \mathbf{W}(\mathbf{I}-\rho\mathbf{W})^{-1}\sigma^2\mathbf{I}\neq\mathbf{0}
\end{aligned}
$$

Εξ' αιτίας της εμφάνισης του προβλήματος της ενδογένειας οι εκτιμητές που θα προκύψουν από την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων δεν θα είναι αμερόληπτοι. Για να αποδειχθεί αυτό, μπορούν να περιληφθούν μαζί σε μια μήτρα **Z** η χωρική υστέρηση **Wy** με τη μήτρα των υπόλοιπων ανεξάρτητων μεταβλητών **X** και σε ένα διάνυσμα **δ** ο συντελεστής ρ με το διάνυσμα των υπόλοιπων συντελεστών **β**, δηλαδή $\mathbf{Z}=\left[\mathbf{Wy}\quad\mathbf{X}\right]$ και $\boldsymbol{\delta}=\begin{bmatrix}\rho\\ \boldsymbol{\beta}\end{bmatrix}$ οπότε το υπόδειγμα (5.5) γράφεται ισοδύναμα ως εξής:

$$\mathbf{y}=\mathbf{Z}\boldsymbol{\delta}+\boldsymbol{\varepsilon} \tag{5.6}$$

και οι εκτιμητές θα δίνονται από τη σχέση $\hat{\boldsymbol{\delta}}=(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$. Η αναμενόμενη τιμή του διανύσματος των εκτιμητών είναι:

$$
\begin{aligned}
\mathrm{E}\left[\hat{\boldsymbol{\delta}}\right] &= \mathrm{E}\left[(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}\right]=\mathrm{E}\left[(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\delta}+\boldsymbol{\varepsilon})\right] \\
&= \mathrm{E}\left[(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Z}\boldsymbol{\delta}+(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}\right] \\
&= \mathrm{E}\left[\boldsymbol{\delta}+(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}\right] \\
&= \mathrm{E}\left[\boldsymbol{\delta}\right]+\mathrm{E}\left[(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}\right] \\
&= \boldsymbol{\delta}+\mathrm{E}\left[(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}\right]
\end{aligned}
$$

Όμως, $\mathrm{E}\left[(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}\right]\neq\mathbf{0}$ διότι όπως έχει αποδειχθεί στη μήτρα **Z** περιέχεται η χωρική υστέρηση **Wy** που συσχετίζεται με το διάνυσμα **ε**. Κατά συνέπεια θα ισχύει ότι:

$$\mathrm{E}\left[\hat{\boldsymbol{\delta}}\right]=\boldsymbol{\delta}+\mathrm{E}\left[\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}\right]\neq\boldsymbol{\delta}$$

και οι εκτιμητές δεν είναι αμερόληπτοι. Αυτό σημαίνει ότι η μέθοδος των ελαχίστων τετραγώνων δεν είναι κατάλληλη για να εκτιμηθεί ένα υπόδειγμα αυτής της μορφής και ο ερευνητής θα πρέπει να επιλέξει κάποια εναλλακτική μέθοδο εκτίμησης όπως η μέθοδος της μέγιστης πιθανοφάνειας ή η μέθοδος των ελαχίστων τετραγώνων σε δυο στάδια.

5.3 Το υπόδειγμα χωρικού σφάλματος

Το υπόδειγμα χωρικού σφάλματος (*Spatial Error Model–SEM*) εκτιμάται όταν ο ερευνητής θεωρεί ότι η παρουσία της αυτοσυσχέτιση στα σφάλματα ενός οικονομετρικού υποδείγματος δεν είναι ουσιαστική (ενόχληση) για την ερμηνεία της εξαρτημένης μεταβλητής αλλά προέρχεται εξαιτίας του είδους των δεδομένων και μπορεί να αντιμετωπιστεί με τον καθορισμό και την εκτίμηση μιας χωρικής διαδικασίας για τα σφάλματα που θα βοηθήσει στον προσδιορισμό των σωστών διακυμάνσεων των εκτιμητών, όπως συμβαίνει και στην επίλυση του προβλήματος της αυτοσυσχέτισης στα οικονομετρικά υποδείγματα όταν χρησιμοποιούνται χρονοσειρές ως δείγμα. Το υπόδειγμα χωρικού σφάλματος με σφάλματα που προέρχονται από μια χωρική αυτοπαλίνδρομη διαδικασία πρώτης τάξης που είναι και η πιο συνηθισμένη περίπτωση ορίζεται ως:

$$y_i=\beta_0+\beta_1 x_{1i}+\ldots+\beta_k x_{ki}+\varepsilon_i \quad \text{με} \quad \varepsilon_i=\lambda\sum_{j=1}^{n} w_{ij}\varepsilon_j+u_i \quad \text{και} \quad u_i\sim iid\,N\left(0,\sigma^2\right)$$

Χρησιμοποιώντας μήτρες το υπόδειγμα γράφεται ως εξής:

$$\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon} \text{ με } \boldsymbol{\varepsilon}=\lambda\mathbf{W}\boldsymbol{\varepsilon}+\mathbf{u}$$

όπου $\mathbf{y}$ είναι το $(n\times 1)$ διάνυσμα της εξαρτημένης μεταβλητής, $\mathbf{X}$ η $(n\times(k+1))$ μήτρα των παρατηρήσεων των k ανεξάρτητων μεταβλητών περιλαμβάνοντας τον σταθερό όρο, $\boldsymbol{\beta}$ το $((k+1)\times 1)$ διάνυσμα των συντελεστών του υποδείγματος, $\mathbf{W}\boldsymbol{\varepsilon}$ είναι η χωρική υστέρηση των σφαλμάτων, λ ο συντελεστής χωρικής υστέρησης και για το διάνυσμα $\mathbf{u}$ γίνεται η υπόθεση ότι $\mathbf{u}\sim N(\mathbf{0},\,\sigma^2\mathbf{I})$. Επιπρόσθετα, για το συντελεστή λ ισχύουν οι συνθήκες ομαλότητας $1/\omega_{\min}<\lambda<1/\omega_{\max}$ όπως παρουσιάστηκαν για τη χωρική αυτοπαλίνδρομη διαδικασία. Ισοδύναμα το υπόδειγμα μπορεί να γραφεί ως:

$$\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\left(\mathbf{I}-\lambda\mathbf{W}\right)^{-1}\mathbf{u}$$

Στο υπόδειγμα χωρικού σφάλματος η ερμηνεία των συντελεστών είναι η ίδια όπως σε ένα απλό οικονομετρικό υπόδειγμα. Η εκτίμηση του υποδείγματος χωρικού σφάλματος επιτυγχάνεται με τη βοήθεια της μεθόδου της μέγιστης πιθανοφάνειας η οποία στη λογαριθμική της μορφή (Anselin, 1988) ορίζεται ως:

$$\ln L\left(\boldsymbol{\beta},\lambda,\sigma^2 \mid \mathbf{y},\mathbf{X}\right) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\sigma^2\right) + \ln\left|\mathbf{I}-\lambda\mathbf{W}\right| - \frac{(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{2\sigma^2}$$

Η λογαριθμική συνάρτηση πιθανοφάνειας συγκεντρώνεται ως προς τις παραμέτρους $\boldsymbol{\beta}$ και σ^2 διευκολύνοντας τη διαδικασία. Ειδικότερα, παραγωγίζοντας ως προς το διάνυσμα $\boldsymbol{\beta}$ προκύπτουν οι εκτιμητές:

$$\hat{\boldsymbol{\beta}}_{ML} = \left[\mathbf{X}'(\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})\mathbf{X}\right]^{-1}\mathbf{X}'(\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})\mathbf{y} = \hat{\boldsymbol{\beta}}_{GLS}$$

οι οποίοι ισούνται (Anselin και Bera, 1998) με τους εκτιμητές $\hat{\boldsymbol{\beta}}_{GLS}$ που προκύπτουν από την εφαρμογή της γενικευμένης μεθόδου των ελαχίστων τετραγώνων (*Generalized Least Squares*), δηλαδή από την παλινδρόμηση με εξαρτημένη μεταβλητή την $(\mathbf{I}-\lambda\mathbf{W})\mathbf{y}$ και ανεξάρτητες τις $(\mathbf{I}-\lambda\mathbf{W})\mathbf{X}$ προϋποθέτοντας ότι η τιμή της παραμέτρου λ είναι γνωστή. Η παραγώγιση ως προς τη διακύμανση σ^2 και η αντικατάσταση των εκτιμηθέντων σφαλμάτων οδηγεί στην εξής εκτίμηση της διακύμανσης:

$$\hat{\sigma}^2_{ML} = \frac{1}{n}\hat{\boldsymbol{\varepsilon}}'_{GLS}(\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})\hat{\boldsymbol{\varepsilon}}_{GLS}$$

Με αντικατάσταση των εκτιμήσεων στη συνάρτηση πιθανοφάνειας προκύπτει η συγκεντρωμένη συνάρτηση πιθανοφάνειας (*concentrated log – likelihood function*), που είναι:

$$\ln L(\lambda) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2} + \ln\left|\mathbf{I}-\lambda\mathbf{W}\right| - \frac{n}{2}\ln\left[\frac{1}{n}\hat{\boldsymbol{\varepsilon}}'_{GLS}(\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})\hat{\boldsymbol{\varepsilon}}_{GLS}\right]$$

η οποία είναι μη γραμμική συνάρτηση ως προς την παράμετρο λ και μπορεί να μεγιστοποιηθεί με την εφαρμογή κάποιας αριθμητικής μεθόδου. Όμως, η διαδικασία είναι περίπλοκη διότι στη συγκεντρωμένη συνάρτηση πιθανοφάνειας του υποδείγματος χωρικού σφάλματος τα εκτιμηθέντα σφάλματα εξαρτώνται έμμεσα από την άγνωστη παράμετρο λ καθώς η εκτίμηση $\hat{\boldsymbol{\beta}}_{GLS}$ απαιτεί τη γνώση της. Για το λόγο αυτό, είναι αναγκαία μια επαναληπτική διαδικασία εκτίμησης που και οι δύο παράμετροι να προσδιορίζονται ταυτόχρονα υπό συνθήκη η μία της άλλης. Ειδικότερα, ο Anselin (1988) προτείνει αρχικά την εκτίμηση με τη μέθοδο των ελαχίστων τετραγώνων του απλού υποδείγματος και τη χρησιμοποίηση των εκτιμηθέντων σφαλμάτων στη συγκεντρωμένη συνάρτηση πιθανοφάνειας ως αρχικά σφάλματα για την εύρεση της τιμής της παραμέτρου λ που τη μεγιστοποιεί.[1] Στη συνέχεια, η τιμή αυτή χρησιμοποιείται για την εύρεση των εκτιμητών $\hat{\boldsymbol{\beta}}_{GLS}$ και τον υπολογισμό νέων σφαλμάτων που αντικαθίστανται στη συγκεντρωμένη συνάρτηση πιθανοφάνειας και οδηγούν σε καινούρια εκτίμηση του λ. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να επέλθει σύγκλιση στα αποτελέσματα σύμφωνα με κάποιο κριτήριο. Η τελική τιμή του λ και τα αντίστοιχα σφάλματα αντικαθίστανται στο τύπο της διακύμανσης για την εκτίμησή της. Αφού βρεθούν όλες οι εκτιμήσεις είναι εφικτό να υπολογιστεί και η μήτρα διακυμάνσεων – συνδιακυμάνσεων των εκτιμητών του υποδείγματος που παρουσιάζεται στους Anselin και Bera (1998). Στη διαδικασία της μεγιστοποίησης διευκολύνει η διάσπαση της ορίζουσας στις ιδιοτιμές της όπως έχει προτείνει ο Ord (1975), δηλαδή ως:

$$\ln\left|\mathbf{I}-\lambda\mathbf{W}\right| = \ln\prod_i\left(1-\lambda\,\omega_i\right) = \sum_i \ln\left(1-\lambda\,\omega_i\right)$$

όπου με ω_i συμβολίζονται οι ιδιοτιμές της μήτρας χωρικών σταθμίσεων $\mathbf{W}$.

5.4 Το υπόδειγμα χωρικής υστέρησης

Το υπόδειγμα χωρικής υστέρησης (*Spatial Lag Model–SLM*) αντιμετωπίζει το πρόβλημα της εμφάνισης χωρικής αυτοσυσχέτισης στα σφάλματα ενός οικονομετρικού υποδείγματος που οφείλεται στην ύπαρξη χωρικών επιδράσεων της μορφής της χωρικής εξάρτησης περιλαμβάνοντας ως ανεξάρτητη μεταβλητή τη χωρική υστέρηση της εξαρτημένης μεταβλητής.[2] Αυτό σημαίνει ότι το πρόβλημα αντιμετωπίζεται με τον ίδιο τρόπο όπως στην ανάλυση χρονοσειρών που προστίθεται η χρονική υστέρηση της εξαρτημένης μεταβλητής στις ανεξάρτητες μεταβλητές όταν υπάρχει χρονική αυτοσυσχέτιση. Η εκτίμηση αυτού του

[1] Ο Ord (1975) περιγράφει μια διαδικασία παρόμοια με τη μέθοδο των Cochrane και Orcutt της ανάλυσης χρονοσειρών στην οποία η τιμή του λ εκτιμάται άμεσα στο SAR(1) υπόδειγμα των σφαλμάτων με τη μέθοδο της μέγιστης πιθανοφάνειας αποφεύγοντας τη μεγιστοποίηση της συγκεντρωμένης πιθανοφάνειας αμφισβητώντας όμως ταυτόχρονα την εγκυρότητά της στη χωρική ανάλυση.

[2] Μερικές φορές το υπόδειγμα χωρικής υστέρησης αναφέρεται και ως χωρικά αυτοπαλίνδρομο υπόδειγμα (*spatial autoregressive model*) ή ως μικτό παλίνδρομο – χωρικά αυτοπαλίνδρομο υπόδειγμα (*mixed regressive – spatial autoregressive model*) εννοώντας ότι πρόκειται για μια χωρικά αυτοπαλίνδρομη διαδικασία στην οποία έχουν προστεθεί ανεξάρτητες μεταβλητές.

υποδείγματος γίνεται όταν η αιτία της χωρικής αυτοσυσχέτισης είναι ουσιαστική, δηλαδή οφείλεται στην αλληλεπίδραση μεταξύ των διαφορετικών γεωγραφικών περιοχών που εκφράζεται μέσω διαδικασιών διάχυσης (*spillover*).

Το υπόδειγμα χωρικής υστέρησης ορίζεται ως:

$$y_i = \rho \sum_{j=1}^{n} w_{ij} y_j + \beta_0 + \beta_1 x_{1i} + \ldots + \beta_k x_{ki} + \varepsilon_i \quad \text{με } \varepsilon_i \sim iid\, N\left(0, \sigma^2\right)$$

Με χρησιμοποίηση άλγεβρας μητρών το υπόδειγμα χωρικής υστέρησης γράφεται ως εξής:

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

όπου $\mathbf{y}$ είναι το ($n \times 1$) διάνυσμα της εξαρτημένης μεταβλητής, $\mathbf{Wy}$ η χωρική υστέρηση της εξαρτημένης μεταβλητής που υπολογίζεται από την ($n \times n$) μήτρα χωρικών σταθμίσεων $\mathbf{W}$, ρ ο συντελεστής χωρικής υστέρησης, $\mathbf{X}$ η ($n \times (k+1)$) μήτρα των παρατηρήσεων των k ανεξάρτητων μεταβλητών περιλαμβάνοντας τον σταθερό όρο, $\boldsymbol{\beta}$ το ($(k+1) \times 1)$) διάνυσμα των συντελεστών του υποδείγματος και $\boldsymbol{\varepsilon}$ το ($n \times 1$) διάνυσμα των τυχαίων σφαλμάτων για τα οποία γίνεται η υπόθεση ότι $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$. Το υπόδειγμα αυτό ισοδύναμα μπορεί να γραφεί ως:

$$\mathbf{y} = \left(\mathbf{I} - \rho \mathbf{W}\right)^{-1} \mathbf{X} \boldsymbol{\beta} + \left(\mathbf{I} - \rho \mathbf{W}\right)^{-1} \boldsymbol{\varepsilon}$$

Από αυτόν τον τρόπο γραφής του υποδείγματος διαπιστώνεται ότι η ερμηνεία των συντελεστών των ανεξάρτητων μεταβλητών δεν πρέπει να γίνεται με τον τρόπο που ισχύει σε ένα απλό οικονομετρικό υπόδειγμα, δηλαδή ως μερικές παράγωγοι που εκφράζουν τη μεταβολή στις τιμές της εξαρτημένης μεταβλητής όταν μεταβάλλεται μία ανεξάρτητη μεταβλητή κατά μία μονάδα ενώ οι υπόλοιπες μεταβλητές του υποδείγματος παραμένουν σταθερές. Ειδικότερα, στο απλό οικονομετρικό υπόδειγμα η μεταβολή μίας ανεξάρτητης μεταβλητής σε μία συγκεκριμένη περιοχή επηρεάζει την τιμή της εξαρτημένης μεταβλητής αποκλειστικά σε αυτή την περιοχή. Αντίθετα, στο υπόδειγμα χωρικής υστέρησης μια μεταβολή στη τιμή κάποιας ανεξάρτητης μεταβλητής σε κάποια περιοχή επηρεάζει τις τιμές της εξαρτημένης μεταβλητής σε όλες τις άλλες περιοχές. Για να αποδειχθεί αυτό, οι LeSage και Pace (2009) βασιζόμενοι στην εργασία των Kim, Phipps και Anselin (2003) γράφουν το υπόδειγμα χωρικής υστέρησης για τις n γεωγραφικές περιοχές (παρατηρήσεις) και τις k ανεξάρτητες μεταβλητές στην εξής μορφή:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{r=1}^{k} \begin{bmatrix} S_r(W)_{11} & S_r(W)_{12} & \cdots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & \cdots & S_r(W)_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ S_r(W)_{n1} & S_r(W)_{n2} & \cdots & S_r(W)_{nn} \end{bmatrix} \begin{bmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{bmatrix} + (\mathbf{I} - \rho\mathbf{W})^{-1} \boldsymbol{\varepsilon}$$

όπου με $S_r(W)_{ij}$ συμβολίζουν το *i, j* στοιχείο στη μήτρα $S_r(W) = (\mathbf{I} - \rho\mathbf{W})^{-1} \beta_r$ όπου β_r ο συντελεστής της *r* ανεξάρτητης μεταβλητής και αν με $V(W)_i$ συμβολιστεί η *i* γραμμή της μήτρας $(\mathbf{I} - \rho\mathbf{W})^{-1}$ τότε το υπόδειγμα χωρική υστέρησης για την *i* περιοχή γράφεται ως:

$$y_i = \sum_{r=1}^{k} \left[S_r(W)_{i1} x_{1r} + S_r(W)_{i2} x_{2r} + \ldots + S_r(W)_{in} x_{nr} \right] + V(W)_i \boldsymbol{\varepsilon}$$

Η μήτρα $S_r(W) = (\mathbf{I} - \rho\mathbf{W})^{-1} \beta_r$ είναι μια μήτρα πολλαπλασιαστής που εφαρμόζει σχέσεις γειτονίας ανωτέρων τάξεων στη μεταβλητή x_r. Εάν συμβολιστούν:

- $S_r(W)_{ij}$ το *ij* στοιχείο της μήτρας $S_r(W)$
- $S_r(W)_{ii}$ το *ii* στοιχείο της μήτρας $S_r(W)$

όπου το *i* αναφέρεται σε μια περιοχή τότε η μερική παράγωγος:

$$\frac{\partial y_i}{\partial x_{ir}} = S_r(W)_{ii} \neq \beta_r \quad \text{(που ισχύει στο απλό γραμμικό υπόδειγμα)}$$

δηλαδή η παράγωγος της *y* στην *i* περιοχή ως προς την x_r μεταβλητή στην ίδια *i* περιοχή θα μετράει την επίδραση των αλλαγών στη μεταβλητή x_r στην *i* περιοχή επάνω στην εξαρτημένη μεταβλητή *y* στην ίδια *i* περιοχή (άμεσες επιδράσεις). Η επίδραση αυτή θα περιλαμβάνει το αποτέλεσμα ανάδρασης όπου η *i* παρατήρηση επηρεάζει την *j* παρατήρηση και η *j* παρατήρηση επηρεάζει την *i* παρατήρηση καθώς και μακρύτερες επιδράσεις από την *i* στην *j* στην *k* και πίσω στην *i* παρατήρηση.

Η παράγωγος της *y* στην *i* περιοχή ως προς την x_r μεταβλητή στην *j* περιοχή

$$\frac{\partial y_i}{\partial x_{jr}} = S_r(W)_{ij} \neq 0 \quad \text{(που ισχύει στο απλό γραμμικό υπόδειγμα)}$$

θα μετράει την επίδραση των αλλαγών στη μεταβλητή x_r στην *j* περιοχή επάνω στην εξαρτημένη μεταβλητή *y* στην *i* περιοχή (έμμεσες επιδράσεις). Στην επίδραση θα περιλαμβάνεται και εδώ αποτέλεσμα ανάδρασης μεταξύ των περιοχών.

Επειδή η ένταση των επιδράσεων διαφέρει από περιοχή σε περιοχή, οι LeSage και Pace (2009) προτείνουν τα παρακάτω τρία στατιστικά μέτρα για να μετριούνται οι χωρικές επιδράσεις.[3]

1) Η μέση άμεση επίδραση

Η μέση άμεση επίδραση (*Average Direct Impact*) υπολογίζεται από τη σχέση:

$$\bar{M}(r)_{direct} = \frac{1}{n} tr\left[S_r\left(W\right)\right]$$

και εκφράζει την επίδραση των μεταβολών στην *i* παρατήρηση της *r* ανεξάρτητης μεταβλητής, x_{ir} επάνω στην *i* παρατήρηση της εξαρτημένης μεταβλητής y_i.

2) Η μέση συνολική επίδραση προς (από) μια παρατήρηση

Η μέση συνολική επίδραση προς μια παρατήρηση (*Average Total Impact to an Observation*) υπολογίζεται από τη σχέση:

$$\bar{M}(r)_{total} = \frac{1}{n} \mathbf{1}' S_r\left(W\right)\mathbf{1}$$

και εκφράζει τη συνολική επίδραση στην *i* παρατήρηση της εξαρτημένης μεταβλητής y_i που οφείλεται από τη μεταβολή της *r* ανεξάρτητης μεταβλητής κατά την ίδια ποσότητα σε όλες τις *n* περιοχές. Η μέση συνολική επίδραση προς μια παρατήρηση θα ισούται και με τη μέση συνολική επίδραση από μια παρατήρηση (*Average Total Impact from an Observation*) που εκφράζει τις συνολικές επιδράσεις σε όλες τις τιμές της *y* που οφείλονται στη μεταβολή της τιμής της *j* παρατήρησης στην *r* ανεξάρτητη μεταβλητή x_{jr}.

3) Η μέση έμμεση επίδραση

Η μέση έμμεση επίδραση (*Average Indirect Impact*) υπολογίζεται από τη σχέση:

$$\bar{M}(r)_{indirect} = \bar{M}(r)_{total} - \bar{M}(r)_{direct}$$

και εκφράζει την έμμεση επίδραση από τη μεταβολή της μεταβλητής x_r.

Η εκτίμηση του υποδείγματος χωρικής υστέρηση δεν μπορεί να γίνει με τη μέθοδο των ελαχίστων τετραγώνων διότι θα προκύψουν μη αμερόληπτοι εκτιμητές λόγω εμφάνισης προβλήματος ενδογένειας καθώς η χωρική υστέρηση της εξαρτημένης μεταβλητής συσχετίζεται με το τυχαίο σφάλμα. Για το λόγο αυτό, για την εκτίμηση του υποδείγματος χρησιμοποιείται η μέθοδος της μέγιστης πιθανοφάνειας. Η λογαριθμική συνάρτηση

[3] Για λεπτομέρειες βλέπε Lesage και Pace (2009) σελίδες 34 – 39.

πιθανοφάνειας για το υπόδειγμα χωρικής υστέρησης σύμφωνα με τον Anselin (1988) ορίζεται ως:

$$\ln L\left(\boldsymbol{\beta},\rho,\sigma^2 \mid \mathbf{y},\mathbf{X}\right)=-\frac{n}{2}\ln\left(2\pi\right)-\frac{n}{2}\ln\left(\sigma^2\right)+\ln\left|\mathbf{I}-\rho\mathbf{W}\right|-\frac{\left(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\right)'\left(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\right)}{2\sigma^2}$$

και πρέπει να μεγιστοποιηθεί ως προς τις παραμέτρους $\boldsymbol{\beta}$, ρ και σ^2. Στη συνάρτηση πιθανοφάνειας υπάρχει και ο λογάριθμος της Ιακωβιανής ορίζουσας $\left|\mathbf{I}-\rho\mathbf{W}\right|$ η οποία θα πρέπει να είναι θετική. Επομένως, θα πρέπει να ισχύουν οι συνθήκες ομαλότητας που σύμφωνα με τον Ord (1975) είναι οι $1/\omega_{\min}<\rho<1/\omega_{\max}$ όπως παρουσιάστηκαν για τη χωρική αυτοπαλίνδρομη διαδικασία.

Η παραγώγιση ως προς το διάνυσμα $\boldsymbol{\beta}$ οδηγεί στους εκτιμητές:

$$\hat{\boldsymbol{\beta}}_{ML}=\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\left(\mathbf{I}-\rho\mathbf{W}\right)\mathbf{y}$$

που εξαρτώνται από την άγνωστη παράμετρο ρ και που ισοδύναμα γράφονται ως:

$$\hat{\boldsymbol{\beta}}_{ML}=\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'-\rho\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}=\hat{\boldsymbol{\beta}}_{ols\mathbf{y},\mathbf{X}}-\rho\hat{\boldsymbol{\beta}}_{ols\mathbf{Wy},\mathbf{X}}$$

Δηλαδή, οι εκτιμητές μέγιστης πιθανοφάνειας διασπώνται σε δύο εκτιμητές ελαχίστων τετραγώνων του $\hat{\boldsymbol{\beta}}_{ols\mathbf{y},\mathbf{X}}$ που προκύπτει από την παλινδρόμηση του υποδείγματος που έχει εξαρτημένη μεταβλητή την $\mathbf{y}$ και ανεξάρτητες μεταβλητές τις $\mathbf{X}$ και του εκτιμητή $\hat{\boldsymbol{\beta}}_{ols\mathbf{Wy},\mathbf{X}}$ που προκύπτει από την παλινδρόμηση του υποδείγματος με εξαρτημένη μεταβλητή τη χωρική υστέρησης $\mathbf{Wy}$ και ανεξάρτητες μεταβλητές τις $\mathbf{X}$. Εφόσον προκύψει με κάποιον τρόπο η εκτίμηση του ρ τότε οι εκτιμητές μέγιστης πιθανοφάνειας υπολογίζονται εύκολα. Για τα εκτιμηθέντα σφάλματα του υποδείγματος ισχύει:

$$\begin{aligned}\hat{\boldsymbol{\varepsilon}}_{ML}&=\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}_{ML}\\&=\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\left(\hat{\boldsymbol{\beta}}_{ols\mathbf{y},\mathbf{X}}-\rho\hat{\boldsymbol{\beta}}_{ols\mathbf{Wy},\mathbf{X}}\right)\\&=\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}_{ols\mathbf{y},\mathbf{X}}+\rho\mathbf{X}\hat{\boldsymbol{\beta}}_{ols\mathbf{Wy},\mathbf{X}}\\&=\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}_{ols\mathbf{y},\mathbf{X}}-\rho\left(\mathbf{W}\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}_{ols\mathbf{Wy},\mathbf{X}}\right)\\&=\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{y},\mathbf{X}}-\rho\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{Wy},\mathbf{X}}\end{aligned}$$

Επομένως και τα σφάλματα από την εκτίμηση με τη μέθοδο της μέγιστης πιθανοφάνειας μπορούν να διασπαστούν στα σφάλματα που προκύπτουν από τις εκτιμήσεις με τη μέθοδο των ελαχίστων τετραγώνων των δύο προαναφερθέντων υποδειγμάτων. Στη συνέχεια παραγωγίζοντας τη συνάρτηση πιθανοφάνειας ως προς τη διακύμανση σ^2 και αντικαθιστώντας τα εκτιμηθέντα σφάλματα προκύπτει:

$$\hat{\sigma}_{ML}^2 = \frac{1}{n}\left(\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{y},\mathbf{X}} - \rho\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{Wy},\mathbf{X}}\right)'\left(\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{y},\mathbf{X}} - \rho\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{Wy},\mathbf{X}}\right)$$

που σημαίνει ότι η διακύμανση από την εκτίμηση του υποδείγματος με τη μέθοδο της μέγιστης πιθανοφάνειας είναι συνάρτηση της άγνωστης παραμέτρου ρ. Κατά συνέπεια, διαπιστώνεται (Anselin, 1988) ότι η λογαριθμική συνάρτηση πιθανοφάνειας μπορεί να συγκεντρωθεί ως προς τις παραμέτρους $\boldsymbol{\beta}$ και σ^2 αντικαθιστώντας τις εκτιμήσεις τους και με τον τρόπο αυτό το πρόβλημα της ταυτόχρονης μεγιστοποίησης πολλών παραμέτρων μετατρέπεται σε πρόβλημα μεγιστοποίησης μίας άγνωστης παραμέτρου της ρ. Η λογαριθμική μορφή της συγκεντρωμένης συνάρτησης πιθανοφάνειας (*concentrated log – likelihood function*) είναι η:

$$\ln L(\rho) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2} + \ln\left|\mathbf{I} - \rho\mathbf{W}\right| - \frac{n}{2}\ln\left[\frac{1}{n}\left(\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{y},\mathbf{X}} - \rho\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{Wy},\mathbf{X}}\right)'\left(\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{y},\mathbf{X}} - \rho\hat{\boldsymbol{\varepsilon}}_{ols\mathbf{Wy},\mathbf{X}}\right)\right]$$

η οποία είναι μη γραμμική συνάρτηση ως προς τη παράμετρο ρ. Η συνάρτηση μεγιστοποιείται με τη βοήθεια κάποιας αριθμητικής μεθόδου και η τιμή του ρ που βρίσκεται αντικαθίσταται στις παραπάνω σχέσεις και υπολογίζονται οι εκτιμητές μέγιστης πιθανοφάνειας των παραμέτρων $\boldsymbol{\beta}$ και σ^2. Στη διαδικασία της μεγιστοποίησης διευκολύνει η διάσπαση της ορίζουσας στις ιδιοτιμές της όπως έχει προτείνει ο Ord (1975), δηλαδή ως:

$$\ln\left|\mathbf{I} - \rho\mathbf{W}\right| = \ln\prod_i\left(1 - \rho\omega_i\right) = \sum_i \ln\left(1 - \rho\omega_i\right)$$

όπου με ω_i συμβολίζονται οι ιδιοτιμές της μήτρας χωρικών σταθμίσεων $\mathbf{W}$. Η μήτρα διακυμάνσεων – συνδιακυμάνσεων των εκτιμητών του υποδείγματος που αναφέρεται στους Anselin και Bera (1998) υπολογίζεται αντιστρέφοντας τη μήτρα πληροφορίας όπως σε όλες τις διαδικασίες εκτίμησης με τη μέθοδο της μέγιστης πιθανοφάνειας.

Παράδειγμα 5.1

Στον Πίνακα 5.1 περιλαμβάνονται οι τιμές για μια μεταβλητής y και μια μεταβλητή x που έχουν καταγραφεί σε πέντε υποθετικές περιοχές. Επίσης δίνεται και η τυποποιημένη ανά γραμμή μήτρα χωρικών σταθμίσεων $\mathbf{W}$ με την οποία καθορίζεται η γειτονία των πέντε χωρικών μονάδων.

Πίνακας 5.1
Τιμές μεταβλητών *y* και *x*

Περιοχή i	y_i	x_i
1	195,77	5
2	215,97	6
3	262,94	16
4	257,41	14
5	265,14	14

$$\mathbf{W} = \begin{bmatrix} 0{,}000 & 1{,}000 & 0{,}000 & 0{,}000 & 0{,}000 \\ 0{,}333 & 0{,}000 & 0{,}333 & 0{,}333 & 0{,}000 \\ 0{,}000 & 0{,}333 & 0{,}000 & 0{,}333 & 0{,}333 \\ 0{,}000 & 0{,}333 & 0{,}333 & 0{,}000 & 0{,}333 \\ 0{,}000 & 0{,}000 & 0{,}500 & 0{,}500 & 0{,}000 \end{bmatrix}$$

Το διάνυσμα τιμών της εξαρτημένης μεταβλητής και η μήτρα των τιμών της ανεξάρτητης μεταβλητής με την προσθήκη μιας στήλης με μονάδες για την εκτίμηση του σταθερού όρου θα είναι, αντίστοιχα:

$$\mathbf{y} = \begin{bmatrix} 195{,}77 \\ 215{,}97 \\ 262{,}94 \\ 257{,}41 \\ 265{,}14 \end{bmatrix} \quad \text{και} \quad \mathbf{X} = \begin{bmatrix} 1 & 5 \\ 1 & 6 \\ 1 & 16 \\ 1 & 14 \\ 1 & 14 \end{bmatrix}$$

Το υπόδειγμα χωρικής υστέρησης εκτιμήθηκε με τη μέθοδο της μέγιστης πιθανοφάνειας και προέκυψε η εκτιμηθείσα σχέση:

$$\hat{y}_i = 0,7153 \sum_{j=1}^{5} w_{ij} \hat{y}_j + 20,205 + 4,207\, x_{1i}$$

δηλαδή:

$$\hat{\rho} = 0{,}7153$$

και

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} 20{,}205 \\ 4{,}207 \end{bmatrix}$$

Για τις ίδιες μεταβλητές εκτιμήθηκε και το απλό οικονομετρικό υπόδειγμα με τη μέθοδο των ελαχίστων τετραγώνων και προέκυψε:

$$\hat{y}_i = 173{,}036 + 6{,}037 x_{1i}$$

Το διάνυσμα των εκτιμηθέντων τιμών των παραμέτρων του απλού υποδείγματος είναι:

$$\hat{\boldsymbol{\beta}}_{OLS} = \begin{bmatrix} 173{,}036 \\ 6{,}037 \end{bmatrix}$$

Η μήτρα $\mathbf{I} - \hat{\rho}\mathbf{W}$ είναι:

$$\mathbf{I} - \hat{\rho}\mathbf{W} = \mathbf{I} - 0{,}7153\mathbf{W} = \begin{bmatrix} 1{,}000 & -0{,}715 & 0{,}000 & 0{,}000 & 0{,}000 \\ -0{,}238 & 1{,}000 & -0{,}238 & -0{,}238 & 0{,}000 \\ 0{,}000 & -0{,}238 & 1{,}000 & -0{,}238 & -0{,}238 \\ 0{,}000 & -0{,}238 & -0{,}238 & 1{,}000 & -0{,}238 \\ 0{,}000 & 0{,}000 & -0{,}358 & -0{,}358 & 1{,}000 \end{bmatrix}$$

Ο χωρικός πολλαπλασιαστής θα είναι:

$$\left(\mathbf{I} - \hat{\rho}\mathbf{W}\right)^{-1} = \left(\mathbf{I} - 0{,}7153\mathbf{W}\right)^{-1} = \begin{bmatrix} 1{,}268 & 1{,}123 & 0{,}453 & 0{,}453 & 0{,}216 \\ 0{,}374 & 1{,}570 & 0{,}633 & 0{,}633 & 0{,}302 \\ 0{,}151 & 0{,}633 & 1{,}505 & 0{,}698 & 0{,}525 \\ 0{,}151 & 0{,}633 & 0{,}698 & 1{,}505 & 0{,}525 \\ 0{,}108 & 0{,}453 & 0{,}788 & 0{,}788 & 1{,}376 \end{bmatrix}$$

Το διάνυσμα εκτιμηθέντων τιμών της *y* για το υπόδειγμα χωρικής υστέρησης είναι:

$$\hat{\mathbf{y}}_{SLM} = \left(\mathbf{I} - \rho\mathbf{W}\right)^{-1} \mathbf{X}\hat{\boldsymbol{\beta}}$$

$$\hat{\mathbf{y}}_{SLM} = \left(\mathbf{I} - 0{,}7153\mathbf{W}\right)^{-1} \mathbf{X}\hat{\boldsymbol{\beta}} = \begin{bmatrix} 1{,}268 & 1{,}123 & 0{,}453 & 0{,}453 & 0{,}216 \\ 0{,}374 & 1{,}570 & 0{,}633 & 0{,}633 & 0{,}302 \\ 0{,}151 & 0{,}633 & 1{,}505 & 0{,}698 & 0{,}525 \\ 0{,}151 & 0{,}633 & 0{,}698 & 1{,}505 & 0{,}525 \\ 0{,}108 & 0{,}453 & 0{,}788 & 0{,}788 & 1{,}376 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 1 & 6 \\ 1 & 16 \\ 1 & 14 \\ 1 & 14 \end{bmatrix} \begin{bmatrix} 20{,}205 \\ 4{,}207 \end{bmatrix} = \begin{bmatrix} 195{,}866 \\ 216{,}171 \\ 263{,}479 \\ 256{,}685 \\ 265{,}139 \end{bmatrix}$$

Το διάνυσμα εκτιμηθέντων τιμών της *y* για το απλό οικονομετρικό υπόδειγμα είναι:

$$\hat{\mathbf{y}}_{OLS} = \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS} = \begin{bmatrix} 1 & 5 \\ 1 & 6 \\ 1 & 16 \\ 1 & 14 \\ 1 & 14 \end{bmatrix} \begin{bmatrix} 173{,}036 \\ 6{,}037 \end{bmatrix} = \begin{bmatrix} 203{,}222 \\ 209{,}259 \\ 269{,}632 \\ 257{,}558 \\ 257{,}558 \end{bmatrix}$$

Ας γίνει η υπόθεση ότι στην περιοχή 3 η τιμή της μεταβλητής x αυξάνεται και από 16 γίνεται 25. Η νέα μήτρα τιμών της ανεξάρτητης μεταβλητής θα είναι:

$$\mathbf{X}^* = \begin{bmatrix} 1 & 5 \\ 1 & 6 \\ 1 & 25 \\ 1 & 14 \\ 1 & 14 \end{bmatrix}$$

Με βάση την αλλαγή στην τιμή της μεταβλητής για την τρίτη περιοχή θα υπολογιστεί των νέο διάνυσμα εκτιμηθέντων τιμών (προβλέψεις) για τα δύο υποδείγματα. Για το απλό οικονομετρικό υπόδειγμα οι προβλεπόμενες τιμές θα είναι:

$$\hat{\mathbf{y}}^*_{OLS} = \mathbf{X}^*\hat{\boldsymbol{\beta}}_{OLS} = \begin{bmatrix} 1 & 5 \\ 1 & 6 \\ 1 & 25 \\ 1 & 14 \\ 1 & 14 \end{bmatrix} \begin{bmatrix} 173{,}036 \\ 6{,}037 \end{bmatrix} = \begin{bmatrix} 203{,}222 \\ 209{,}259 \\ 323{,}968 \\ 257{,}558 \\ 257{,}558 \end{bmatrix}$$

Όπως διαπιστώνεται οι εκτιμηθείσες τιμές παραμένουν οι ίδιες για τις περιοχές 1,2,4 και 5 και αλλάζει μόνο η προβλεπόμενη τιμή στην περιοχή 3 (αύξηση από 269,632 σε 323,968), δηλαδή στην περιοχή που έχει γίνει η μεταβολή στην τιμή της ανεξάρτητης μεταβλητής. Για το υπόδειγμα χωρικής υστέρησης οι καινούριες προβλεπόμενες τιμές θα είναι:

$$\hat{\mathbf{y}}^*_{SLM} = \left(\mathbf{I} - 0{,}7153\mathbf{W}\right)^{-1}\mathbf{X}\hat{\boldsymbol{\beta}} = \begin{bmatrix} 1{,}268 & 1{,}123 & 0{,}453 & 0{,}453 & 0{,}216 \\ 0{,}374 & 1{,}570 & 0{,}633 & 0{,}633 & 0{,}302 \\ 0{,}151 & 0{,}633 & 1{,}505 & 0{,}698 & 0{,}525 \\ 0{,}151 & 0{,}633 & 0{,}698 & 1{,}505 & 0{,}525 \\ 0{,}108 & 0{,}453 & 0{,}788 & 0{,}788 & 1{,}376 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 1 & 6 \\ 1 & 25 \\ 1 & 14 \\ 1 & 14 \end{bmatrix} \begin{bmatrix} 20{,}205 \\ 4{,}207 \end{bmatrix} = \begin{bmatrix} 213{,}016 \\ 240{,}147 \\ 320{,}472 \\ 283{,}103 \\ 294{,}970 \end{bmatrix}$$

Όπως διαπιστώνεται, οι εκτιμηθείσες τιμές έχουν αλλάξει και στις πέντε περιοχές και έχουν γίνει μεγαλύτερες. Αυτό οφείλεται στα αποτελέσματα διάχυσης που προκαλούνται από το χωρικό πολλαπλασιαστή.

Για τη μεταβλητή x η μήτρα $S_r(W) = (\mathbf{I} - \rho\mathbf{W})^{-1}\ \beta_r$ θα είναι (r = 1):

$$S_1(W) = (\mathbf{I} - 0{,}7153\mathbf{W})^{-1}\ \hat{\beta}_1 = (\mathbf{I} - 0{,}7153\mathbf{W})^{-1}\ 4{,}207 = \begin{bmatrix} 5{,}333 & 4{,}724 & 1{,}906 & 1{,}906 & 0{,}909 \\ 1{,}575 & 6{,}604 & 2{,}664 & 2{,}664 & 1{,}270 \\ 0{,}635 & 2{,}664 & 6{,}332 & 2{,}935 & 2{,}210 \\ 0{,}635 & 2{,}664 & 2{,}935 & 6{,}332 & 2{,}210 \\ 0{,}454 & 1{,}906 & 3{,}315 & 3{,}315 & 5{,}788 \end{bmatrix}$$

Δηλαδή πολλαπλασιάζεται η μήτρα του χωρικού πολλαπλασιαστή με την εκτιμηθείσα τιμή της παραμέτρου β_1 από το υπόδειγμα χωρικής υστέρησης.

Η μέση άμεση επίδραση θα είναι:

$$\bar{M}(r)_{direct} = \frac{1}{n} tr\left[S_1(W)\right] = \frac{30{,}38983}{5} = 6{,}077965$$

και στην πραγματικότητα είναι ο μέσος όρος των στοιχείων της κυρίας διαγωνίου της μήτρας $S_1(W)$.

Η μέση συνολική επίδραση προς μια παρατήρηση θα είναι:

$$\bar{M}(r)_{total} = \frac{1}{n}\mathbf{1}'S_1(W)\mathbf{1} = \frac{73{,}884}{5} = 14{,}77681$$

και στην πραγματικότητα πρόκειται για το μέσο όρο του αθροίσματος των γραμμών (ή στηλών) της μήτρας $S_1(W)$.

Η μέση έμμεση επίδραση θα είναι:

$$\bar{M}(r)_{indirect} = \bar{M}(r)_{total} - \bar{M}(r)_{direct} = 14{,}77681 - 6{,}077965 = 8{,}698842$$

5.5 Έλεγχοι για την ύπαρξη χωρικής αυτοσυσχέτισης στα σφάλματα του γραμμικού υποδείγματος παλινδρόμησης

Στην ενότητα αυτή παρουσιάζονται οι πιο σημαντικοί έλεγχοι που συνήθως χρησιμοποιούνται στη χωρική οικονομετρία για τον έλεγχο της ύπαρξης χωρικής αυτοσυσχέτισης στα σφάλματα ενός απλού οικονομετρικού υποδείγματος και τον εντοπισμό της αιτίας δημιουργίας του προβλήματος. Πιο συγκεκριμένα, περιγράφονται η διαδικασία εφαρμογής του ελέγχου με τη στατιστικής I του Moran, που αναφέρθηκε στο Κεφάλαιο 3, για τα σφάλματα ενός γραμμικού υποδείγματος παλινδρόμησης και μια κατηγορία διαγνωστικών ελέγχων που συμβάλλουν στον εντοπισμό της γενεσιουργού αιτίας του προβλήματος και στην εξειδίκευση του σωστού χωρικού οικονομετρικού υποδείγματος.

5.5.1 Ο έλεγχος *I* του Moran για σφάλματα παλινδρόμησης

Η πιο γνωστή διαδικασία έλεγχου για την ύπαρξη χωρικής αυτοσυσχέτισης στα σφάλματα ενός γραμμικού υποδείγματος παλινδρόμησης είναι με την εφαρμογή της στατιστικής I του Moran (1950) όπως έχει προταθεί από τους Cliff και Ord (1972) για την περίπτωση καταλοίπων παλινδρόμησης. Η στατιστική I του Moran ορίζεται ως:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\hat{\boldsymbol{\varepsilon}}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}$$

όπου n είναι το μέγεθος του δείγματος. $\hat{\boldsymbol{\varepsilon}}$ το $(n \times 1)$ διάνυσμα των καταλοίπων που έχουν προκύψει από την εκτίμηση του υποδείγματος με τη μέθοδο των ελαχίστων τετραγώνων, $\mathbf{W}$ είναι η $(n \times n)$ μήτρα χωρικών σταθμίσεων και $\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$ ένας παράγοντας κανονικοποίησης που ισούται με το άθροισμα όλων των χωρικών σταθμίσεων. Επειδή για μια τυποποιημένη κατά γραμμή μήτρα χωρικών σταθμίσεων ισχύει $\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} = n$ η στατιστική I του Moran γράφεται στη μορφή:

$$I = \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\hat{\boldsymbol{\varepsilon}}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}$$

Οι Cliff και Ord (1972) απέδειξαν την ασυμπτωτική κατανομή της στατιστικής I όταν η εξεταζόμενη μεταβλητή αποτελείται από κατάλοιπα παλινδρόμησης. Ειδικότερα, υπό τη μηδενική υπόθεση της απουσίας χωρικής αυτοσχέτισης και υποθέτοντας κανονικότητα η στατιστική *I* του Moran έχει μέση τιμή και διακύμανση που ορίζονται αντίστοιχα ως:

$$\mathrm{E}(I) = \frac{tr\left[\mathbf{MW}\right]}{n-(k+1)}$$

και

$$\mathrm{Var}(I) = \frac{tr\left[\mathbf{MWMW'}\right] + tr\left[\mathbf{MW}\right]^2 + \left(tr\left[\mathbf{MW}\right]\right)^2}{(n-k-1)(n-k+1)} - \left[\mathrm{E}(I)\right]^2,$$

όπου $\mathbf{M} = \mathbf{I} - \mathbf{X}\left(\mathbf{X'X}\right)^{-1}\mathbf{X'}$ και k είναι ο αριθμός των ανεξάρτητων μεταβλητών του υποδείγματος. Κατά συνέπεια, εφόσον το δείγμα είναι αρκετά μεγάλο, η μετασχηματισμένη μεταβλητή:

$$Z = \frac{I - \mathrm{E}(I)}{\sqrt{\mathrm{Var}(I)}}$$

ακολουθεί την τυποποιημένη κανονική κατανομή και μπορεί να χρησιμοποιηθεί για τη διεξαγωγή του ελέγχου της στατιστικής σημαντικότητας του *I*. Η εφαρμογή του ελέγχου που βασίζεται στην κατασκευή, με διαδικασία μεταθέσεων, της εμπειρικής κατανομή της *I* υπό τη μηδενική υπόθεση ή που βασίζεται στην υπόθεση της τυχαιοποίησης οδηγεί σε αναξιόπιστα αποτελέσματα στην περίπτωση των καταλοίπων παλινδρόμησης.

Αξίζει να αναφερθεί ότι ενώ ως μηδενική υπόθεση του ελέγχου ορίζεται η μη ύπαρξη χωρικής αυτοσυσχέτισης δεν υπάρχει κάποια συγκεκριμένη μορφή για την εναλλακτική υπόθεση και απλά υποτίθεται η παρουσία κάποιου τύπου χωρικής αυτοσυσχέτισης. Στην πράξη, έχει αποδειχθεί ότι ο έλεγχος αυτός έχει ισχύ για αρκετές μορφές χωρικής αυτοσυσχέτισης με πιο συνηθισμένες τη χωρική αυτοπαλίνδρομη διαδικασία και τη χωρική διαδικασία κινητού μέσου. Επιπρόσθετα, η γνωστή από την ανάλυση χρονοσειρών στατιστική των Durbin και Watson αποτελεί ειδική περίπτωση της στατιστικής *I* όταν η μήτρα χωρικών σταθμίσεων οριστεί με ειδικό τρόπο.[4] Ωστόσο, η εφαρμογή της στατιστικής *I* του Moran για χωρικά δεδομένα παρουσιάζει πρόβλημα καθορισμού κριτικών τιμών, κάτι το οποίο δεν ισχύει με το στατιστικό έλεγχο των Durbin και Watson. Το πρόβλημα αυτό επικεντρώνεται στη δυσκολία των μαθηματικών υπολογισμών που απαιτούνται και στο ότι οι προκύπτουσες κριτικές τιμές, με τη μεθοδολογία που έχουν αναπτύξει οι Tiefelsdorf και Boots (1995), εξαρτώνται από τη μήτρα χωρικών σταθμίσεων η οποία διαφέρει σε κάθε

[4] Για λεπτομέρειες βλέπε Anselin και Bera (1998) σελίδα 266.

χωρική ανάλυση. Για το λόγο αυτό, ο έλεγχος εφαρμόζεται με κριτικές τιμές από την τυπική κανονική κατανομή, όπως έχουν δείξει οι Cliff και Ord (1972).

5.5.2 Οι έλεγχοι χωρικής εξάρτησης

Ο έλεγχος *I* του Moran αποτελεί σημαντικό διαγνωστικό εργαλείο για τον εντοπισμό της ύπαρξης χωρικής αυτοσυσχέτισης στα σφάλματα ενός γραμμικού υποδείγματος αλλά δεν παρέχει καμία πληροφορία για τον τύπο της αυτοσυσχέτισης και την αιτία που την προκαλεί. Όπως έχει αναφερθεί, η αυτοσυσχέτιση στα σφάλματα μπορεί να είναι αποτέλεσμα χωρικής αυτοσυσχέτισης στην εξαρτημένη μεταβλητή αλλά και κάποιου μη μετρήσιμου παράγοντα που επηρεάζει την εξαρτημένη μεταβλητή. Και στις δύο περιπτώσεις ο ερευνητής θα πρέπει να εκτιμήσει κάποιο υπόδειγμα που περιλαμβάνει χωρικές επιδράσεις. Τα δύο πιο συχνά χρησιμοποιούμενα χωρικά υποδείγματα είναι το υπόδειγμα χωρικής υστέρησης (*Spatial Lag Model*) και το υπόδειγμα χωρικού σφάλματος (*Spatial Error Model*). Στην ενότητα αυτή παρουσιάζονται μια κατηγορία στατιστικών ελέγχων που συνεισφέρουν στον εντοπισμό της αιτίας της εμφάνισης χωρικής αυτοσυσχέτισης και στην εξειδίκευση του σωστού χωρικού οικονομετρικού υποδείγματος. Οι έλεγχοι αυτοί είναι γνωστοί ως έλεγχοι χωρικής εξάρτησης (*Spatial Dependence Tests*) και ανήκουν στην κατηγορία των ελέγχων σκορ ή πολλαπλασιαστή Lagrange (*Score test* ή *Lagrange Multiplier test*). Το βασικό τους πλεονέκτημα είναι ότι ο υπολογισμός των στατιστικών τους βασίζεται αποκλειστικά στα κατάλοιπα που προκύπτουν από την εκτίμηση του απλού οικονομετρικού υποδείγματος, με τη μέθοδο των ελαχίστων τετραγώνων, χωρίς να απαιτείται η εκτίμηση του αντίστοιχου χωρικού οικονομετρικού υποδείγματος. Αξίζει να αναφερθεί ότι σε αντίθεση με τους περισσότερους ελέγχους πολλαπλασιαστή Lagrange που εφαρμόζονται στην οικονομετρία, η διαδικασία εφαρμογής των έλεγχων χωρικής εξάρτησης δεν περιλαμβάνει το στάδιο υπολογισμού του συντελεστή προσδιορισμού από κάποια βοηθητική παλινδρόμηση. Η περιγραφή των ελέγχων βασίζεται στο γενικό χωρικό οικονομετρικό υπόδειγμα του Anselin (1988), γνωστό και με την ονομασία SAC (LeSage και Pace, 2009).[5]

Ειδικότερα, το γενικό χωρικό οικονομετρικό υπόδειγμα, το οποίο παρουσιάζεται λεπτομερώς στην εργασία των Anselin, Bera, Florax, και Yoon (1996), αποτελείται από ένα

[5] O Elhorst (2014) στη σελίδα 6 αναφέρει ότι δεν υπάρχει κάποια συγκεκριμένη ερμηνεία του συμβολισμού SAC. Προφανώς, ο συμβολισμός αναφέρεται στα αρχικά των ονομάτων LeSage και Pace (2009).

ταυτόχρονο σύστημα δύο εξισώσεων που μπορούν να ορίσουν το απλό οικονομετρικό υπόδειγμα, το υπόδειγμα χωρικής υστέρησης (*Spatial Lag Model – SLM*) και το υπόδειγμα χωρικού σφάλματος (*Spatial Error Model – SEM*) και έχει ως εξής:

$$\begin{aligned} &\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ &\boldsymbol{\varepsilon} = \lambda \mathbf{W}_2 \boldsymbol{\varepsilon} + \mathbf{u} \\ &\mathbf{u} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \end{aligned} \tag{5.7}$$

όπου $\mathbf{y}$ είναι το ($n \times 1$) διάνυσμα της εξαρτημένης μεταβλητής, $\mathbf{X}$ η ($n \times (k+1)$) μήτρα των k ανεξάρτητων μεταβλητών περιλαμβάνοντας μία στήλη με μονάδες για την εκτίμηση του σταθερού όρου, $\boldsymbol{\beta}$ το ($(k+1) \times 1$) διάνυσμα των παραμέτρων, $\boldsymbol{\varepsilon}$ και $\mathbf{u}$ τα ($n \times 1$) διανύσματα των σφαλμάτων και ρ και λ οι συντελεστές χωρικών υστερήσεων της εξαρτημένης μεταβλητής και των σφαλμάτων. Οι ($n \times n$) μήτρες $\mathbf{W}_1$ και $\mathbf{W}_2$ περιέχουν τις χωρικές σταθμίσεις για τον ορισμό των χωρικών υστερήσεων στις δύο εξισώσεις του υποδείγματος. Για λόγους απλοποίησης θεωρείται ότι $\mathbf{W}_1 = \mathbf{W}_2 = \mathbf{W}$ και ότι στη μήτρα $\mathbf{X}$ υπάρχει μία τουλάχιστον ανεξάρτητη μεταβλητή εκτός του σταθερού όρου ώστε να μπορεί το υπόδειγμα (5.7) να ταυτοποιηθεί διότι διαφορετικά μπορεί να εμφανιστούν προβλήματα προσδιορισμού των παραμέτρων ρ και λ. Είναι φανερό ότι εάν στο υπόδειγμα (5.7) τεθούν $\rho = \lambda = 0$ προκύπτει το απλό οικονομετρικό υπόδειγμα, εάν τεθεί $\lambda = 0$ προκύπτει το υπόδειγμα χωρικής υστέρησης και εάν τεθεί $\rho = 0$ το υπόδειγμα χωρικού σφάλματος.

Ο πρώτος έλεγχος αυτής της κατηγορίας εξετάζει την αναγκαιότητα εκτίμησης του υποδείγματος χωρικού σφάλματος σε σχέση με το απλό υπόδειγμα και παρουσιάστηκε αρχικά από τον Burridge (1980). Δηλαδή, υποθέτοντας ότι $\rho = 0$ η μηδενική και η εναλλακτική υπόθεση ορίζονται αντίστοιχα ως:

$$\begin{aligned} &H_0 : \lambda = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα} \\ &H_1 : \lambda \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικού Σφάλματος} \end{aligned}$$

Για να εκτελεστεί ο έλεγχος εκτιμάται με τη μέθοδο των ελαχίστων τετραγώνων το απλό οικονομετρικό υπόδειγμα και υπολογίζεται η στατιστική συνάρτηση:

$$\text{LM-ERR} = \frac{\left[\dfrac{\hat{\boldsymbol{\varepsilon}}' \mathbf{W} \hat{\boldsymbol{\varepsilon}}}{S^2} \right]^2}{T} \tag{5.8}$$

η οποία υπό τη μηδενική υπόθεση ακολουθεί ασυμπτωτικά την κατανομή χ^2 με ένα βαθμό ελευθερίας. Στη σχέση (5.8) με $\hat{\boldsymbol{\varepsilon}}$ συμβολίζεται το διάνυσμα των καταλοίπων από την εκτίμηση, $S^2 = \hat{\boldsymbol{\varepsilon}}' \hat{\boldsymbol{\varepsilon}} / n$ είναι η εκτιμηθείσα διακύμανση της παλινδρόμησης και

$T = tr\left[\mathbf{W}^2 + \mathbf{W}'\mathbf{W}\right]$. Αποδεικνύεται ότι η στατιστική LM-ERR είναι η ίδια και για την περίπτωση που η εναλλακτική υπόθεση του ελέγχου είναι μια χωρική διαδικασία κινητού μέσου της μορφής $\boldsymbol{\varepsilon} = \lambda\mathbf{W}\mathbf{u} + \mathbf{u}$ κάτι που συνεπάγεται ότι πρακτικά είναι αδύνατη η διάκριση μεταξύ των δύο τύπων αυτοσυσχέτισης. Επιπρόσθετα, η στατιστική LM-ERR μπορεί ισοδύναμα να γραφεί ως:

$$\text{LM-ERR} = \frac{\left[n\dfrac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\hat{\boldsymbol{\varepsilon}}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}\right]^2}{T}$$

από όπου προκύπτει ότι η στατιστική αυτή στην πραγματικότητα είναι το τετράγωνο της στατιστικής I του Moran διαιρεμένο με τον παράγοντα T/n^2 κάτι που εξηγεί και την ισχύ της για διαφορετικές μορφές χωρικής αυτοσυσχέτισης.

Ο έλεγχος πολλαπλασιαστή Lagrange για την παρουσία στο υπόδειγμα χωρικής υστέρησης στην εξαρτημένη μεταβλητή προτάθηκε από τον Anselin (1988b) και υποθέτοντας ότι $\lambda = 0$ έχει ως μηδενική και εναλλακτική υπόθεση:

$$\begin{aligned} H_0 &: \rho = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα} \\ H_1 &: \rho \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικής Υστέρησης} \end{aligned}$$

και πραγματοποιείται με βάση την ακόλουθη στατιστική:

$$\text{LM-LAG} = \frac{\left[\dfrac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\mathbf{y}}{S^2}\right]^2}{T + \left[\dfrac{\left(\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}}\right)'\mathbf{M}\left(\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}}\right)}{S^2}\right]} \tag{5.9}$$

όπου $\hat{\boldsymbol{\beta}}$ είναι το διάνυσμα με τις εκτιμηθείσες τιμές των παραμέτρων του απλού υποδείγματος και $\mathbf{M} = \mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'$. Η στατιστική LM-LAG έχει την ίδια ασυμπτωτική κατανομή με την LM-ERR, δηλαδή ακολουθεί την κατανομή χ^2 με ένα βαθμό ελευθερίας.

Οι στατιστικές LM-ERR και LM-LAG δεν είναι έγκυρες όταν αντίστοιχα $\rho \neq 0$ και $\lambda \neq 0$. Οι Anselin, Bera, Florax, και Yoon (1996) ανέπτυξαν ανθεκτικές (*robust*) μορφές των ελέγχων που διορθώνουν για την ύπαρξη τοπικής λανθασμένης εξειδίκευσης (*local misspecification*) και μπορούν να χρησιμοποιηθούν για την επιλογή του σωστού χωρικού οικονομετρικού υποδείγματος όταν υπάρχει ασάφεια από την εφαρμογή των απλών ελέγχων, σημειώνοντας ότι ως τοπική λανθασμένη εξειδίκευση εννοείται ότι οι τιμές των λ και ρ δεν είναι μηδέν αλλά πολύ κοντά στο μηδέν. Ωστόσο, πρέπει να αναφερθεί πως και οι ανθεκτικές μορφές των ελέγχων αντιμετωπίζουν προβλήματα όταν οι τιμές των λ και ρ είναι πολύ μακριά από το μηδέν. Ο ανθεκτικός έλεγχος της υπόθεσης $H_0 : \lambda = 0$ που λαμβάνει υπ' όψιν την παρουσία χωρικής υστέρησης της εξαρτημένης μεταβλητής πραγματοποιείται με τη στατιστική:

$$\text{LM-EL} = \frac{\left[\frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\hat{\boldsymbol{\varepsilon}}}{S^2} - T \frac{S^2}{S^2 T + \left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right)' \mathbf{M} \left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right)} \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\mathbf{y}}{S^2} \right]^2}{T - T^2 \left[\frac{S^2}{S^2 T + \left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right)' \mathbf{M} \left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right)} \right]} \qquad (5.10)$$

και για τον έλεγχο της υπόθεσης $H_0 : \rho = 0$ που λαμβάνει υπ' όψιν την παρουσία χωρικής αυτοσυσχέτισης στα σφάλματα με τη στατιστική:

$$\text{LM-LE} = \frac{\left[\frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\mathbf{y}}{S^2} - \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\hat{\boldsymbol{\varepsilon}}}{S^2} \right]^2}{\frac{1}{S^2}\left[\left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right)' \mathbf{M} \left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right) + S^2 T \right] - T} \qquad (5.11)$$

Οι στατιστικές αυτές όπως και οι μη ανθεκτικές εκδόσεις τους ακολουθούν την κατανομή χ^2 με ένα βαθμό ελευθερίας.

Ένας άλλος τρόπος διενέργειας των ελέγχων με τον οποίο δεν απαιτείται η εξασφάλιση ότι η μία από τις δύο χωρικές παραμέτρους λαμβάνει την τιμή μηδέν είναι να πραγματοποιηθεί από κοινού έλεγχος ότι υπάρχει χωρική υστέρηση στην εξαρτημένη μεταβλητή και ταυτόχρονα τα σφάλματα ακολουθούν μια χωρική αυτοπαλίνδρομη διαδικασία ή μια διαδικασία κινητού μέσου. Ο έλεγχος αυτός που παρουσιάστηκε από τον Anselin (1988b) ορίζεται ως:

$$H_0 : \lambda = \rho = 0$$
$$H_1 : \text{Ένα τουλάχιστον απο τα } \lambda, \rho \neq 0$$

και γίνεται με τη στατιστική:

$$\text{LM-SARMA} = \frac{\left[\frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\hat{\boldsymbol{\varepsilon}}}{S^2} \right]^2}{T} + \frac{\left[\frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\mathbf{y}}{S^2} - \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{W}\hat{\boldsymbol{\varepsilon}}}{S^2} \right]^2}{\frac{1}{S^2}\left[\left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right)' \mathbf{M} \left(\mathbf{WX}\hat{\boldsymbol{\beta}}\right) + S^2 T \right] - T} \qquad (5.12)$$

η οποία υπολογίζεται από την εκτίμηση του απλού υποδείγματος και υπό τη μηδενική υπόθεση ακολουθεί την κατανομή χ^2 με δύο βαθμούς ελευθερίας. Από τη σχέση (5.12) διαπιστώνεται ότι LM-SARMA = LM-ERR + LM-LE. Ανάλογα, μετά από μερικές αλγεβρικές πράξεις, προκύπτει ότι LM-SARMA = LM-LAG + LM-EL. Κατά συνέπεια, η στατιστική LM-SARMA αποτελείται από το άθροισμα της στατιστικής του απλού ελέγχου μίας χωρικής παραμέτρου με τη στατιστική του ανθεκτικού ελέγχου της άλλης παραμέτρου.

Ο από κοινού έλεγχος για την ύπαρξη των δύο τύπων χωρικής εξάρτησης έχει δύο βασικά μειονεκτήματα. Το πρώτο είναι ότι όταν απορρίπτεται η μηδενική υπόθεση ο ερευνητής δεν έχει καμία ένδειξη για το χωρικό οικονομετρικό υπόδειγμα που θα πρέπει να εκτιμήσει και το δεύτερο είναι ότι ο έλεγχος έχει απώλεια σε ισχύ όταν στην πραγματικότητα υπό τη μηδενική υπόθεση εμφανίζεται μόνο η μία μορφή χωρικής εξάρτησης.

Οι έλεγχοι χωρικής εξάρτησης που παρουσιάστηκαν μερικές φορές οδηγούν σε αντιφατικά συμπεράσματα ως προς την επιλογή του σωστού χωρικού οικονομετρικού υποδείγματος. Στις περιπτώσεις που υπάρχει ασάφεια από τα αποτελέσματα των ελέγχων οι Anselin και Florax (1995) και οι Anselin, Bera, Florax, και Yoon (1996) προτείνουν την ακόλουθη στρατηγική εντοπισμού του σωστού υποδείγματος. Αρχικά διεξάγονται οι έλεγχοι με τις στατιστικές LM-ERR και LM-LAG. Εάν κανένας από τους δύο ελέγχους δεν απορρίπτει τη μηδενική του υπόθεση τότε πρέπει να εξεταστούν τα κατάλοιπα του απλού οικονομετρικού υποδείγματος για την πιθανότητα ύπαρξης άλλων προβλημάτων που μπορεί να οδήγησαν στην εσφαλμένη απόρριψη του ελέγχου χωρικής αυτοσυσχέτισης με τη στατιστική I του Moran όπως η ετεροσκεδαστικότητα και η έλλειψη κανονικότητας. Εάν στον έναν από τους δύο ελέγχους απορρίπτεται η μηδενική υπόθεση, τότε επιλέγεται και εκτιμάται το αντίστοιχο χωρικό υπόδειγμα. Στην περίπτωση που απορρίπτεται η μηδενική υπόθεση και στους δύο απλούς ελέγχους διεξάγονται οι έλεγχοι με τις στατιστικές LM-EL και LM-LE και επιλέγεται το χωρικό υπόδειγμα του ελέγχου που απορρίπτεται η μηδενική του υπόθεση. Εάν και στους δύο ανθεκτικούς ελέγχους απορρίπτεται η μηδενική υπόθεση τότε επιλέγεται και εκτιμάται το υπόδειγμα που υποδεικνύει η ανθεκτική στατιστική με τη μεγαλύτερη τιμή.

Παράδειγμα 5.2

Ένα απλό οικονομετρικό υπόδειγμα με δυο ανεξάρτητες μεταβλητές και δείγμα n = 100 παρατηρήσεων εκτιμήθηκε με τη μέθοδο των ελαχίστων τετράγωνων. Η στατιστική *I* του Moran για τα σφάλματα του υποδείγματος μαζί με τη μέση τιμή της, τη διακύμανσή της, τη μετασχηματισμένη μεταβλητή *Z* και το P-Value για τον έλεγχο της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης δίνονται στον Πίνακα 5.2. Αντίστοιχα, στον Πίνακα 5.3 δίνονται οι τιμές και τα P-Values για τους LM ελέγχους χωρικής εξάρτησης.

Πίνακας 5.2
Στατιστική *I* του Moran για τα εκτιμημένα σφάλματα του παραδείγματος 5.2

I	E(*I*)	Var(*I*)	Z-Value	P-Value
-0,02375	-0,00966	0,005464	-0,1906	0,8488

Πίνακας 5.3
Τιμές LM στατιστικών χωρικής εξάρτησης του παραδείγματος 5.2

Στατιστική LM	Τιμή	P-Value
LM-ERR	0,0991	0,7529
LM-LAG	0,0293	0,8641
LM-EL	0,0740	0,7856
LM-LE	0,0042	0,9482

Αρχικά πραγματοποιείται ο έλεγχος χωρικής αυτοσυσχέτισης για τα σφάλματα του υποδείγματος. Ο έλεγχος της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης γίνεται με τη στατιστική:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1)$$

η οποία με βάση τις πληροφορίες στον Πίνακα 5.2 θα είναι:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} = \frac{-0,02375 - (-0,00966)}{\sqrt{0,005464}} = -0,1906$$

Κατά συνέπεια, για επίπεδο σημαντικότητας α = 5% η κριτική τιμή θα είναι $Z_{0,025}$ = 1,96 οπότε επειδή $|Z| = 0,19061 < 1,96$ η μηδενική υπόθεση του ελέγχου δεν απορρίπτεται και επομένως αφού δεν υπάρχουν ενδείξεις χωρική αυτοσυσχέτισης δεν έχει νόημα η επιλογή ενός χωρικού υποδείγματος με τη βοήθεια των LM ελέγχων χωρικής εξάρτησης.

Παράδειγμα 5.3

Ένα απλό οικονομετρικό υπόδειγμα με δυο ανεξάρτητες μεταβλητές και δείγμα n = 100 παρατηρήσεων εκτιμήθηκε με τη μέθοδο των ελαχίστων τετράγωνων. Η στατιστική *I* του Moran για τα σφάλματα του υποδείγματος μαζί με τη μέση τιμή της, τη διακύμανσή της, τη μετασχηματισμένη μεταβλητή *Z* και το P-Value για τον έλεγχο της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης δίνονται στον Πίνακα 5.4. Αντίστοιχα, στον Πίνακα 5.5 περιλαμβάνονται οι τιμές και τα P-Values για τους LM ελέγχους χωρικής εξάρτησης.

Πίνακας 5.4
Στατιστική *I* του Moran για τα εκτιμημένα σφάλματα του παραδείγματος 5.3

I	E(*I*)	Var(*I*)	Z-Value	P-Value
0,70081	-0,0100639	0,00552052	9,5677	2,2e-16

Πίνακας 5.5
Τιμές LM στατιστικών χωρικής εξάρτησης του παραδείγματος 5.3

Στατιστική LM	Τιμή	P-Value
LM-ERR	86,3342	2,2e-16
LM-LAG	34,9375	3,405e-09
LM-EL	52,7234	3,84e-13
LM-LE	1,3266	0,2494

Αρχικά πραγματοποιείται ο έλεγχος χωρικής αυτοσυσχέτισης για τα σφάλματα του υποδείγματος. Ο έλεγχος της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης γίνεται με τη στατιστική:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1)$$

η οποία με βάση τις πληροφορίες στον Πίνακα 5.4 θα είναι:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} = \frac{0,70081 - (-0,0100639)}{\sqrt{0,00552052}} = 9,5677$$

Κατά συνέπεια, για επίπεδο σημαντικότητας α = 5% η κριτική τιμή θα είναι $Z_{0,025}$ = 1,96 οπότε επειδή $|Z| = 9,5677 > 1,96$ η μηδενική υπόθεση του ελέγχου απορρίπτεται και επομένως αφού υπάρχουν ενδείξεις χωρική αυτοσυσχέτισης θα χρησιμοποιηθούν οι LM

έλεγχοι χωρικής εξάρτησης από τον Πίνακα 5.5 για την επιλογή ενός χωρικού οικονομετρικού υποδείγματος.

- για τον LM έλεγχο χωρικού σφάλματος (υπόθεση ότι ρ = 0) προκύπτει:

$$H_0 : \lambda = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \lambda \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικού Σφάλματος}$$

Επειδή LM-ERR = 86,3342 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

- για τον LM έλεγχο χωρικής υστέρησης (υπόθεση ότι λ = 0) προκύπτει:

$$H_0 : \rho = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \rho \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικής Υστέρησης}$$

Επειδή LM-LAG = 34,9375 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

Αφού και οι δυο έλεγχοι με τις απλές LM στατιστικές απέρριψαν τις μηδενικές τους υποθέσεις θα γίνουν οι έλεγχοι με τις ανθεκτικές στατιστικές.

- για τον ανθεκτικό LM έλεγχο χωρικού σφάλματος (υπόθεση ότι $\rho \approx 0$) προκύπτει

$$H_0 : \lambda = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \lambda \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικού Σφάλματος}$$

Επειδή LM-EL = 52,7234 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

- για τον ανθεκτικό LM έλεγχο χωρικής υστέρησης (υπόθεση ότι $\lambda \approx 0$) προκύπτει:

$$H_0 : \rho = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \rho \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικής Υστέρησης}$$

Επειδή LM-LE = 1,3266 < $X^2_{1,0,05}$ = 3,84 δεν απορρίπτεται η μηδενική υπόθεση.

Επομένως , εφόσον ο ανθεκτικός LM έλεγχος χωρικού σφάλματος απορρίπτει τη μηδενική του υπόθεση και ο ανθεκτικός LM έλεγχος χωρικής υστέρησης δεν την απορρίπτει θα πρέπει να επιλεχθεί και να εκτιμηθεί το υπόδειγμα χωρικού σφάλματος.

Παράδειγμα 5.4

Ένα απλό οικονομετρικό υπόδειγμα με δυο ανεξάρτητες μεταβλητές και δείγμα n = 100 παρατηρήσεων εκτιμήθηκε με τη μέθοδο των ελαχίστων τετράγωνων. Η στατιστική *I* του Moran για τα σφάλματα του υποδείγματος μαζί με τη μέση τιμή της, τη διακύμανσή της, τη μετασχηματισμένη μεταβλητή *Z* και το P-Value για τον έλεγχο της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης δίνονται στον Πίνακα 5.6. Αντίστοιχα, στον Πίνακα 5.7 περιλαμβάνονται οι τιμές και τα P-Values για τους LM ελέγχους χωρικής εξάρτησης.

Πίνακας 5.6
Στατιστική *I* του Moran για τα εκτιμημένα σφάλματα του παραδείγματος 5.4

I	E(*I*)	Var(*I*)	Z-Value	P-Value
0,620223	-0,009955	0,0054859	8,5083	2,2e-16

Πίνακας 5.7
Τιμές LM στατιστικών χωρικής εξάρτησης του παραδείγματος 5.4

Στατιστική LM	Τιμή	P-Value
LM-ERR	67,6197	2,22e-16
LM-LAG	125,8258	2,2e-16
LM-EL	0,1563	0,6925
LM-LE	125,9821	2,176e-14

Αρχικά πραγματοποιείται ο έλεγχος χωρικής αυτοσυσχέτισης για τα σφάλματα του υποδείγματος. Ο έλεγχος της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης γίνεται με τη στατιστική:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1)$$

η οποία με βάση τις πληροφορίες στον Πίνακα 5.6 θα είναι:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} = \frac{0,620223 - (-0,009955)}{\sqrt{0,0054859}} = 8,5083$$

Κατά συνέπεια, για επίπεδο σημαντικότητας α = 5% η κριτική τιμή θα είναι $Z_{0,025}$ = 1,96 οπότε επειδή $|Z| = 8,5083 > 1,96$ η μηδενική υπόθεση του ελέγχου απορρίπτεται και επομένως αφού υπάρχουν ενδείξεις χωρική αυτοσυσχέτισης θα χρησιμοποιηθούν οι LM

έλεγχοι χωρικής εξάρτησης από τον Πίνακα 5.7 για την επιλογή ενός χωρικού οικονομετρικού υποδείγματος.

- για τον LM έλεγχο χωρικού σφάλματος (υπόθεση ότι $\rho = 0$) προκύπτει:

$$H_0 : \lambda = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \lambda \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικού Σφάλματος}$$

Επειδή LM-ERR = 67,6197 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

- για τον LM έλεγχο χωρικής υστέρησης (υπόθεση ότι $\lambda = 0$) προκύπτει:

$$H_0 : \rho = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \rho \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικής Υστέρησης}$$

Επειδή LM-LAG = 125,8258 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

Αφού και οι δυο έλεγχοι με τις απλές LM στατιστικές απέρριψαν τις μηδενικές τους υποθέσεις θα γίνουν οι έλεγχοι με τις ανθεκτικές στατιστικές.

- για τον ανθεκτικό LM έλεγχο χωρικού σφάλματος (υπόθεση ότι $\rho \approx 0$) προκύπτει

$$H_0 : \lambda = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \lambda \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικού Σφάλματος}$$

Επειδή LM-EL = 0,1563 < $X^2_{1,0,05}$ = 3,84 δεν απορρίπτεται η μηδενική υπόθεση.

- για τον ανθεκτικό LM έλεγχο χωρικής υστέρησης (υπόθεση ότι $\lambda \approx 0$) προκύπτει:

$$H_0 : \rho = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \rho \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικής Υστέρησης}$$

Επειδή LM-LE = 125,9821 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

Επομένως , εφόσον ο ανθεκτικός LM έλεγχος χωρικής υστέρησης απορρίπτει τη μηδενική του υπόθεση και ο ανθεκτικός LM έλεγχος χωρικού σφάλματος δεν την απορρίπτει θα πρέπει να επιλεχθεί και να εκτιμηθεί το υπόδειγμα χωρικής υστέρησης.

Παράδειγμα 5.5

Ένα απλό οικονομετρικό υπόδειγμα με δυο ανεξάρτητες μεταβλητές και δείγμα n = 100 παρατηρήσεων εκτιμήθηκε με τη μέθοδο των ελαχίστων τετράγωνων. Η στατιστική *I* του Moran για τα σφάλματα του υποδείγματος μαζί με τη μέση τιμή της, τη διακύμανσή της, τη μετασχηματισμένη μεταβλητή Z και το P-Value για τον έλεγχο της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης δίνονται στον Πίνακα 5.8. Αντίστοιχα, στον Πίνακα 5.9 περιλαμβάνονται οι τιμές και τα P-Values για τους LM ελέγχους χωρικής εξάρτησης.

Πίνακας 5.8
Στατιστική *I* του Moran για τα εκτιμημένα σφάλματα του παραδείγματος 5.5

I	E(*I*)	Var(*I*)	Z-Value	P-Value
0,718996	-0,009612	0,002883	13,5693	2,2e-16

Πίνακας 5.9
Τιμές LM στατιστικών χωρικής εξάρτησης του παραδείγματος 5.5

Στατιστική LM	Τιμή	P-Value
LM-ERR	168,632	2,2e-16
LM-LAG	207,9474	2,2e-16
LM-EL	11,6371	0,0006465
LM-LE	50,9525	9,462e-13

Αρχικά πραγματοποιείται ο έλεγχος χωρικής αυτοσυσχέτισης για τα σφάλματα του υποδείγματος. Ο έλεγχος της μηδενικής υπόθεσης της απουσίας χωρικής αυτοσυσχέτισης γίνεται με τη στατιστική:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1)$$

η οποία με βάση τις πληροφορίες στον Πίνακα 5.8 θα είναι:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} = \frac{0,718996 - (-0,009612)}{\sqrt{0,002883}} = 13,5693$$

Κατά συνέπεια, για επίπεδο σημαντικότητας α = 5% η κριτική τιμή θα είναι $Z_{0,025}$ = 1,96 οπότε επειδή $|Z| = 13,5693 > 1,96$ η μηδενική υπόθεση του ελέγχου απορρίπτεται και επομένως αφού υπάρχουν ενδείξεις χωρική αυτοσυσχέτισης θα χρησιμοποιηθούν οι LM

έλεγχοι χωρικής εξάρτησης από τον Πίνακα 5.9 για την επιλογή ενός χωρικού οικονομετρικού υποδείγματος.

- για τον LM έλεγχο χωρικού σφάλματος (υπόθεση ότι ρ = 0) προκύπτει:

$$H_0 : \lambda = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \lambda \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικού Σφάλματος}$$

Επειδή LM-ERR = 168,632 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

- για τον LM έλεγχο χωρικής υστέρησης (υπόθεση ότι λ = 0) προκύπτει:

$$H_0 : \rho = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \rho \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικής Υστέρησης}$$

Επειδή LM-LAG = 207,9474 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

Αφού και οι δυο έλεγχοι με τις απλές LM στατιστικές απέρριψαν τις μηδενικές τους υποθέσεις θα γίνουν οι έλεγχοι με τις ανθεκτικές στατιστικές.

- για τον ανθεκτικό LM έλεγχο χωρικού σφάλματος (υπόθεση ότι $\rho \approx 0$) προκύπτει

$$H_0 : \lambda = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \lambda \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικού Σφάλματος}$$

Επειδή LM-EL = 11,6371 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

- για τον ανθεκτικό LM έλεγχο χωρικής υστέρησης (υπόθεση ότι $\lambda \approx 0$) προκύπτει:

$$H_0 : \rho = 0 \quad \Rightarrow \text{ Απλό Οικονομετρικό Υπόδειγμα}$$
$$H_1 : \rho \neq 0 \quad \Rightarrow \text{ Υπόδειγμα Χωρικής Υστέρησης}$$

Επειδή LM-LE = 50,9525 > $X^2_{1,0,05}$ = 3,84 απορρίπτεται η μηδενική υπόθεση.

Εφόσον τόσο ο ανθεκτικός LM έλεγχος χωρικού σφάλματος όσο και ο ανθεκτικός LM έλεγχος χωρικής υστέρησης απορρίπτουν τις μηδενικές τους υποθέσεις ο ερευνητής βρίσκεται σε δίλλημα σχετικά με το ποιο χωρικό υπόδειγμα πρέπει να επιλέξει. Σε αυτή την περίπτωση, σύμφωνα με εργασίες προσομοίωσης που έχουν παρουσιαστεί στη βιβλιογραφία, πρέπει να επιλεχθεί και να εκτιμηθεί το υπόδειγμα που υποδεικνύει η ανθεκτική στατιστική με τη μεγαλύτερη τιμή. Κατά συνέπεια, επειδή LM-LE = 50,9525 > LM-EL = 11,6371 θα πρέπει να εκτιμηθεί το υπόδειγμα χωρικής υστέρησης. Αξίζει να παρατηρηθεί ότι το ίδιο αποτέλεσμα προκύπτει και εάν γινόταν η σύγκριση μεταξύ των απλών LM στατιστικών.