Supplementary Material for: Can cross-lagged panel modeling be relied on to establish cross-lagged effects? The case of contemporaneous and reciprocal effects

Bengt Muthén and Tihomir Asparouhov

November 28, 2023

1 Appendix

1.1 Identification proof for RCLPM

Consider the CLPM for the variables Y_t and Z_t for $t = 2, ..., T$,

$$Y_t = \alpha_{yt} + \beta_{1t} Y_{t-1} + \beta_{2t} Z_{t-1} + \varepsilon_{yt} \tag{1}$$

$$Z_t = \alpha_{zt} + \beta_{3t} Y_{t-1} + \beta_{4t} Z_{t-1} + \varepsilon_{zt} \tag{2}$$

$$\varepsilon_{yt} \sim N(0, v_{yt}) \tag{3}$$

$$\varepsilon_{zt} \sim N(0, v_{zt}) \tag{4}$$

$$c_t = Cov(\varepsilon_{yt}, \varepsilon_{zt}). \tag{5}$$

Next, consider the reciprocal cross-lagged model.This model is referred to as RCLPM (reciprocal cross-lagged panel model). The RCLPM can be expressed as

$$Y_t = a_{yt} + r_{yt} Z_t + b_{1t} Y_{t-1} + b_{2t} Z_{t-1} + \varepsilon_{yt} \tag{6}$$

$$Z_t = a_{zt} + r_{zt} Y_t + b_{3t} Y_{t-1} + b_{4t} Z_{t-1} + \varepsilon_{zt} \tag{7}$$

$$\varepsilon_{yt} \sim N(0, w_{yt}) \tag{8}$$

$$\varepsilon_{zt} \sim N(0, w_{zt}) \tag{9}$$

$$0 = Cov(\varepsilon_{yt}, \varepsilon_{zt}). \tag{10}$$

We begin with the case of $T = 3$. In this case we need $2 = T - 1$ parameter constraints to make the RCLPM model identifiable and possibly equivalent to the CLPM. We consider the following constraint: time invariance of the reciprocal interactions

$$r_{y2} = r_{y3} = r_y \tag{11}$$

$$r_{z2} = r_{z3} = r_z. \tag{12}$$

Under this constraint, we have the following 6 equations, obtained from equations (**??**-**??**) for $t = 2$ and $t = 3$:

$$v_{y2} = \frac{w_{y2} + r_y^2 w_{z2}}{(1 - r_y r_z)^2} \tag{13}$$

$$v_{z2} = \frac{w_{z2} + r_z^2 w_{y2}}{(1 - r_y r_z)^2} \tag{14}$$

$$c_2 = \frac{w_{z2} r_y + w_{y2} r_z}{(1 - r_y r_z)^2} \tag{15}$$

$$v_{y3} = \frac{w_{y3} + r_y^2 w_{z3}}{(1 - r_y r_z)^2} \tag{16}$$

$$v_{z3} = \frac{w_{z3} + r_z^2 w_{y3}}{(1 - r_y r_z)^2} \tag{17}$$

$$c_3 = \frac{w_{z3} r_y + w_{y3} r_z}{(1 - r_y r_z)^2} \tag{18}$$

In these 6 equations, the CLPM has 6 parameters $v_{y2}, v_{z2}, c_2, v_{y3}, v_{z3}, c_3$ and the RCLPM has 6 parameters $w_{y2}, w_{z2}, w_{y3}, w_{z3}, r_y, r_z$. The above 6 equations show how to derive the CLPM parameters from the RCLPM parameters. If we can reverse these equations and show that the 6 RCLPM parameters can be derived from the 6 CLPM parameters, we establish the equivalence of the two models under the constraints (11-12). Thus we need to solve the above equations for the RCLPM parameters. Below we delve into this task. Using (13) and (15),

$$w_{y2} = (v_{y2} - c_2 r_y)(1 - r_y r_z). \tag{19}$$

Similarly,

$$w_{z2} = (v_{z2} - c_2 r_z)(1 - r_y r_z) \tag{20}$$

$$w_{y3} = (v_{y3} - c_3 r_y)(1 - r_y r_z) \tag{21}$$

$$w_{z3} = (v_{z3} - c_3 r_z)(1 - r_y r_z). \tag{22}$$

Thus, if we find a way to solve for r_y and r_z in terms of the CLPM parameters, the above 4 equations will complete the task. Substituting w_{y2} and w_{z2}, using (19) and (20), in equation (15) we obtain

$$v_{z2} r_y + v_{y2} r_z = (1 + r_y r_z) c_2 \tag{23}$$

and similarly

$$v_{z3} r_y + v_{y3} r_z = (1 + r_y r_z) c_3. \tag{24}$$

These two equations can serve as the basis for determining r_y and r_z in terms of the CLPM parameters. Next we divide the two equations to obtain

$$\frac{v_{z2}}{c_2} r_y + \frac{v_{y2}}{c_2} r_z = \frac{v_{z3}}{c_3} r_y + \frac{v_{y3}}{c_3} r_z \tag{25}$$

and

$$r_z = r_y \frac{\frac{v_{z2}}{c_2} - \frac{v_{z3}}{c_3}}{\frac{v_{y3}}{c_3} - \frac{v_{y2}}{c_2}}. \tag{26}$$

With this last equation we are solving for r_z and then only r_y is left, however, we see that a new condition for identification appears

$$\frac{v_{y3}}{c_3} \neq \frac{v_{y2}}{c_2} \tag{27}$$

and

$$\frac{v_{z3}}{c_3} \neq \frac{v_{z2}}{c_2}. \tag{28}$$

This can be interpreted as follows. There has to be some time non-invariance in v_{yt}/c_t and v_{zt}/c_t for the identification to occur. Also note that in the situation when that non-invariance is not very pronounced, the identification of the reciprocal regression parameters is likely to be poor. Furthermore, for finite sample size when the non-invariance is not sufficiently pronounced to ensure that the asymptotic distribution of the denominator in (26) is away from zero, we can expect a very non-normal/skewed parameter distribution for the estimated reciprocal regression parameters. Such a distribution can be assessed properly with the Bayes estimator or with the bootstrap method, both of which allow for asymmetric parameter distribution. The ML estimator, which assumes a symmetric asymptotic distribution, may yield questionable confidence intervals that do not reflect the skewed distribution.

The final step in this analysis is to determine an expression for r_y in terms of the CLPM parameters. Denote the quantity

$$\lambda = \frac{\frac{v_{z2}}{c_2} - \frac{v_{z3}}{c_3}}{\frac{v_{y3}}{c_3} - \frac{v_{y2}}{c_2}}, \tag{29}$$

so that

$$r_z = r_y \lambda. \tag{30}$$

We then substitute that expression in equation (23) and obtain the following quadratic equation

$$\lambda r_y^2 - \Big(\frac{v_{z2}}{c_2} + \lambda \frac{v_{y2}}{c_2}\Big) r_y + 1 = 0 \tag{31}$$

with discriminant

$$D = \Big(\frac{v_{z2}}{c_2} + \lambda \frac{v_{y2}}{c_2}\Big)^2 - 4\lambda.$$

Using the basic inequality that $(a+b)^2 \geq 4ab$ and the fact that $v_{y2} v_{z2} \geq c_2^2$, we can see that $D \geq 0$ and therefore the quadratic equation always yields a solution

$$r_y = \frac{\frac{v_{z2}}{c_2} + \lambda \frac{v_{y2}}{c_2} \pm \sqrt{D}}{2\lambda}. \tag{32}$$

As is typical in quadratic equations we get two different solutions. That may present somewhat of an interpretation challenge because there is no statistical way to discriminate between the two. Furthermore, in simulation studies, there is no guarantee that different replications will converge towards the same solution. It may be that some of the replications converge to one solution and some to the other, rendering the typical Montecarlo evaluation strategies useless. The two solutions can also cause problems for the Bayes and bootstrap estimators because the built parameter distribution may become a bimodal mixture of the posterior/bootstrap distribution for the two solutions. Inequality constraints on the reciprocal regression parameters can be used to reduce the posterior/bootstrap distribution to only one of the two solutions. If the sample size is small, however, separating the posterior/bootstrap distributions for the two solutions may become impossible because the two distributions will overlap substantially.

1.2 Resolving the dual solution problem

In this section we provide some more information on the Appendix Section 1.1 dual solution problem that arises from the quadratic nature of the equations determining the reciprocal regression parameters of the RCLPM. Suppose that the two solutions of equation (31) are r'_y and r''_y and the corresponding solutions for r_z are r'_z and r''_z. It can be shown that

$$\frac{r'_z}{r'_y} = \frac{r''_z}{r''_y} = \lambda \tag{33}$$

$$r'_y r''_y = \frac{1}{\lambda} \tag{34}$$

$$r'_z r''_z = \lambda \tag{35}$$

$$r'_y r'_z r''_y r''_z = 1, \tag{36}$$

where λ is given in (29). Because of equation (36), in one of the two solutions $|r_y r_z| < 1$ and in the other $|r_y r_z| > 1$. This inequality can be used to assist Monte Carlo simulations, Bayes and bootstrap estimations to converge to just one of the two solutions. For example, constraining the estimation to the case of

$$r_y^2 r_z^2 < 1 \tag{37}$$

will ensure that all estimates across Monte Carlo replications, MCMC draws, or bootstrap draws are using just one of the two solutions.

In principle the two solutions are mathematically equivalent, however, we can argue here that the solution which satisfies $|r_y r_z| < 1$ will be easier to interpret. The product of the two reciprocal regression coefficients represents the feedback loop and one would generally expect that to be less than 1. Furthermore, there is the impact on the auto-regressive (AR) matrix. In the typical stationary CLPM, both eigenvalues as well as the determinant of the AR matrix will be between 0 and 1. In the RCLPM, that determinant is multiplied by $1 - r_y r_z$ which will be negative if $r_y r_z > 1$, i.e., the RCLPM will have an unusual AR matrix. Finally, the solution which satisfies $|r_y r_z| < 1$ has the advantage that the structural concepts of total and indirect effects are actually defined, see Chapter 8 in Bollen (1989). This is because the eigenvalues of

$$B_0 = \begin{pmatrix} 0 & r_y \\ r_z & 0 \end{pmatrix}$$

are less than 1 by absolute value precisely when $|r_y r_z| < 1$. When these eigenvalues are less than 1 by absolute value, B_0^n converges to zero and that guarantees that the total and indirect effects can be computed for the reciprocal model. The matrix B_0^n contains the indirect paths of length n.

In principle, the dual solution problem can be eliminated with additional constraints on the model parameters. For example, in the article Section The RCLPM with time-invariant reciprocal and cross-lagged regression we show that if the cross-lagged regressions are invariant across time, there is no dual solution. In the article Section The RCLPM without cross-lagged regressions and Appendix Section 1.5, we show that if the cross-lagged regressions are fixed to zero, there is no dual solution. Not every parameter constraint however can eliminate the dual solution. Consider as an example the case of $T > 3$ where all reciprocal regression coefficients are time invariant. This means that we have $T - 1$ equations to determine r_y and r_z of the type given in (23-24) instead of just two such equations. Unfortunately, there can be only two independent such equations. Any other equation of the same type will be a linear combination of the first two. To be more specific, suppose that we have the following system of three equations

$$p_1 r_y + q_1 r_z = 1 + r_y r_z$$

$$p_2 r_y + q_2 r_z = 1 + r_y r_z$$

$$p_3 r_y + q_3 r_z = 1 + r_y r_z.$$

Subtracting the third equation from the first two yields

$$(p1 - p3)r_y = (q3 - q1)r_z$$

$$(p2 - p3)r_y = (q3 - q2)r_z.$$

If r_y and r_z are non-zero, we can divide these two equations and we obtain

$$\frac{p1 - p3}{p2 - p3} = \frac{q1 - q3}{q2 - q3}.$$

If we denote the above quantity by δ then

$$p_3 = \frac{1}{1-\delta} p_1 - \frac{\delta}{1-\delta} p_2$$

$$q_3 = \frac{1}{1-\delta} q_1 - \frac{\delta}{1-\delta} q_2,$$

which means that the third equation is a linear combination of the first two and it does not carry any new information for r_y and r_z. We conclude that reciprocal interaction invariance constraint for $T > 3$ will not provide any additional information that can resolve the quadratic nature of the solution we have for $T = 3$.

1.3 Resolving interpretability problems due to negative R^2

When an RCLPM is estimated, a negative R^2 value can occur for some of the variables. In this section we discuss conditions for when that occurs and common sense strategies to deal with this problem. Mplus computes the following R^2 values for the RCLPM (6-7),

$$R^2_{yt} = 1 - \frac{Var(\varepsilon_{yt})}{Var(Y_t)}$$

$$R^2_{zt} = 1 - \frac{Var(\varepsilon_{zt})}{Var(Z_t)}.$$

These quantities are somewhat intractable because they involve auto-regressive and reciprocal relationships. Here we consider the conditional R^2 values, where we condition on all variables from the previous period

$$R^2_{yt0} = 1 - \frac{Var(\varepsilon_{yt})}{Var(Y_t | Y_{t-1}, Z_{t-1})}$$

$$R^2_{zt0} = 1 - \frac{Var(\varepsilon_{zt})}{Var(Z_t | Y_{t-1}, Z_{t-1})}.$$

The conditional R^2 values can also be viewed as the unconditional R^2 if all auto-regressive parameters are 0. It can also be viewed as the incremental improvements in R^2 obtained by the predictors we don't condition on. All of the above R^2 values should be positive. If any of these are negative, interpretability will be compromised. Negative R^2 values imply the illogical conclusion that when a predictor is added to a regression equation, the error becomes bigger rather than smaller. When we add a predictor to a regression, we expect the predictor to help explain the variation in the predicted variable and to reduce the residual error term. As we add predictors we expect the R^2 to increase monotonically. In this section we show that this expectation fails precisely when r_y and r_z are of opposite signs.

In what follows we focus on the RCLPM for $T = 3$ with invariant reciprocal parameters. However, the conclusions can be extended to other models. We consider the RCLPM solution where $|r_y r_z| < 1$, which we established earlier as the most interpretable case. We will show that if $0 < r_y r_z < 1$ the conditional R^2 are positive and if $-1 < r_y r_z < 0$ at least one of the conditional R^2 is negative. First note that

$$R^2_{yt0} = 1 - \frac{w_{yt}}{v_{yt}}$$

where v_{yt} is given in (**??**). Therefore

$$R^2_{yt0} = \frac{w_{yt} + r_{yt}^2 w_{zt} - w_{yt}(1 - r_{yt} r_{zt})^2}{w_{yt} + r_{yt}^2 w_{zt}} = \frac{w_{yt} r_{yt} r_{zt}(2 - r_{yt} r_{zt}) + r_{yt}^2 w_{zt}}{w_{yt} + r_{yt}^2 w_{zt}}$$

which is clearly positive when $0 < r_y r_z < 1$.

Now consider the case $-1 < r_y r_z < 0$. The reciprocal parameters are of opposite sign. One of the two reciprocals will have a sign opposite to the sign of c_2. Let's assume that is r_y, i.e., $r_y c_2 < 0$. From equation (20) we then see that $w_{yt} > v_{yt}$ which implies that R_{yt0} is negative.

An alternative argument that implies difficulties with the interpretation of the case $r_y r_z < 0$ goes as follows. If we substitute equation (7) in (6) we obtain an equation

$$Y_t = ... + r_y r_z Y_t + ...$$

If $r_y r_z < 0$ the equation implies that an increase in Y_t leads to a decrease in Y_t which is a contradiction.

Next we focus on the condition of the CLPM parameters that determine whether or not the equivalent RCLPM will be interpretable, i.e., r_y and r_z would have the same sign. Because of equation (30), we see that r_y and r_z have the same sign if and only if $\lambda > 0$. Furthermore we can see from equation (29) that if c_2 and c_3 are of opposite signs then $\lambda < 0$ and the RCLPM would not be interpretable. Thus, if the sign of the residual covariance of the CLPM changes over time, the RCLPM is not interpretable. It is also possible to show that $\lambda > 0$ is equivalent to the following condition

$$\max(\overline{\rho_t}, 1/\overline{\rho_t}) < \max(\overline{v_{yt}}/\overline{v_{zt}}, \overline{v_{zt}}/\overline{v_{yt}}),$$

where $\overline{\rho_t} = \rho_t/\rho_{t-1}$ is the rate of change in the residual correlation parameter ρ_t of the VAR model, and $\overline{v_{yt}} = \sqrt{v_{yt}/v_{yt-1}}$ is the rate of change in the standard deviation of ε_{yt}, etc. The above inequality can be interpreted as follows. In order for a CLPM to produce an interpretable RCLPM, the correlation parameter in the CLPM must have more stability across time than the ratio of the scales of the residuals.

The R^2 issue applies to any reciprocal and more generally non-recursive SEM models, i.e., it applies to the RCLPM discussed above but also the models in article Section 3.2 and Section 1.5 below. In certain situations, it may be preferable to deal with the negative R^2 problem not by adjusting the model but by adjusting the definition of the R^2, see Hayduk (2006). This approach should be reserved for those situations when a very strong substantive reasoning is available (in favor of a nonrecursive model) that outweighs the opposing evidence found in the data (the negative R^2 should be interpreted as evidence against the model). In almost all situations, minor modifications of a nonrecursive model can convert it to a recursive model, preserving model fit and resolving the negative R^2 issue. Such a modification, for example, can be replacing a regression parameter with a covariance parameter. Non-recursive models have an abundance of competing/alternative recursive models which will not have a negative R^2. It would be difficult to argue in general that all of these recursive alternative models should be dismissed. This is particularly the case for the RCLPM which has a perfectly reasonable and well established alternative CLPM.

1.4 The RCLPM with invariant reciprocal and cross-lagged regressions, and non-invariant residual covariances

In this section we show that the RCLPM with invariant reciprocal and cross-lagged regressions with added residual covariances is an identified model as long as the auto-regressive parameters b_{1t} and b_{4t} are not time invariant. To do that, we first consider a constrained version of the CLPM (1-5). The constraint that we are interested in can be described as follows: the scatter plot of the Y_t regression coefficients (β_{1t}, β_{3t}) forms a straight line, and the scatter plot of the Z_t regression coefficients (β_{4t}, β_{2t}) forms a straight line as well.[1] This amounts to adding the following parameter constraints

$$\frac{\beta_{3,t+1} - \beta_{3t}}{\beta_{1,t+1} - \beta_{1t}} = \frac{\beta_{33} - \beta_{32}}{\beta_{13} - \beta_{12}} \tag{38}$$

$$\frac{\beta_{2,t+1} - \beta_{2t}}{\beta_{4,t+1} - \beta_{4t}} = \frac{\beta_{23} - \beta_{22}}{\beta_{43} - \beta_{42}} \tag{39}$$

or an equivalent version of those. These equations can easily be converted to an expression which specifies that β_{3t} is a linear function of β_{1t} and β_{2t} is a linear function of β_{4t}. Since the CLPM is identified, so is the constrained CLPM with the constraints (38-39). Let's now denote the slope and intercept of the scatter plot line of (β_{1t}, β_{3t}) by r_z and b_3, and denote the slope and intercept of the scatter plot line of (β_{4t}, β_{2t}) by r_y and b_2. The expression given in (38) is r_z and the expression given in (39) is r_y. Note that as long as the scatter plot for (β_{1t}, β_{3t}) contains at least two distinct points, i.e., β_{1t} is not time invariant, the slope and intercept for the straight line r_z and b_3 are identified. Similarly, as long as β_{4t} is not time invariant, r_y and b_2 are identified. Because of the linearity we obtain

$$\beta_{3t} = r_z \beta_{1t} + b_3 \tag{40}$$

$$\beta_{2t} = r_y \beta_{4t} + b_2. \tag{41}$$

[1] Adding such constraints in Mplus can be accomplished with the MODEL CONSTRAINT command.

Next we define b_{1t} and b_{4t} as follows

$$b_{1t} = \beta_{1t} - r_y\beta_{3t} \tag{42}$$

$$b_{4t} = \beta_{4t} - r_z\beta_{2t}. \tag{43}$$

Now it is easy to see that equations (40-43) are equivalent to article equation (23) under the assumption of invariant reciprocal RVAR parameters $r_{yt} = r_y$, $r_{zt} = r_z$ and invariant cross-lagged parameters $b_{2t} = b_2$ and $b_{3t} = b_3$. We conclude that the constrained CLPM, using constraints (38-39), is a reparameterization of the RCLPM with invariant reciprocal and cross-lagged parameters. Note that the auto-regressive parameters b_{1t} and b_{4t} are not time invariant precisely when β_{1t} and β_{4t} are not time invariant. This is the only condition needed for the above reparameterization.

The reparametrization for the remaining model parameters (intercepts, variances and covariances) is given as follows. The reparameterization for the intercept parameters is given again by article equation (17) or equivalently by equation (22). The reparameterization for the residual covariance parameters is derived as follows. If ε_{yt} and ε_{zt} are the residuals of the VAR model, from article equation (13) we see that the residuals for the RCLPM are

$$\begin{pmatrix} 1 & -r_y \\ -r_z & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix}. \tag{44}$$

If we denote the residual covariance parameter of the RCLPM by ρ_t, the variance/covariance reparameterization is given by

$$\begin{pmatrix} w_{yt} & \rho_t \\ \rho_t & w_{zt} \end{pmatrix} = \begin{pmatrix} 1 & -r_y \\ -r_z & 1 \end{pmatrix} \begin{pmatrix} v_{yt} & c_t \\ c_t & v_{zt} \end{pmatrix} \begin{pmatrix} 1 & -r_z \\ -r_y & 1 \end{pmatrix} \tag{45}$$

or equivalently

$$\begin{pmatrix} v_{yt} & c_t \\ c_t & v_{zt} \end{pmatrix} = \frac{1}{(1 - r_y r_z)^2} \begin{pmatrix} 1 & r_y \\ r_z & 1 \end{pmatrix} \begin{pmatrix} w_{yt} & \rho_t \\ \rho_t & w_{zt} \end{pmatrix} \begin{pmatrix} 1 & r_z \\ r_y & 1 \end{pmatrix}. \tag{46}$$

We conclude that the RCLPM with invariant reciprocal and cross-lagged regressions and added residual covariances is identified because it is equivalent to a constrained RCLPM. The above derivation also implies that this RCLPM does not have a dual solution but it can be a subject to the negative R^2 issue discussed in Appendix Section 1.3. As in Appendix Section 1.3, the conditional R^2 for the RVAR model is positive if and only if

$$v_{yt} > w_{yt} \tag{47}$$

$$v_{zt} > w_{zt}. \tag{48}$$

From (45) we get

$$v_{yt} > w_{yt} = v_{yt} - 2r_y c_t + r_y^2 v_{zt}$$

$$v_{zt} > w_{zt} = v_{zt} - 2r_z c_t + r_z^2 v_{yt}$$

or equivalently

$$2r_y c_t > r_y^2 v_{zt}$$

$$2r_z c_t > r_z^2 v_{yt}$$

If we multiply the above two inequalities we obtain that for the RCLPM to have positive conditional R^2 it is necessary to have $r_y r_z > 0$, i.e., the reciprocal coefficients must have the same sign. If they have different signs, the model has a negative conditional R^2 and thus the model is inadmissible. This conclusion is the same as in Appendix Section 1.3 for the RCLPM without the residual covariance. Multiplying the above two inequalities also gives as a necessary upper bound for $r_y r_z$

$$r_y r_z < \frac{4c_t^2}{v_{yt} v_{zt}} < 4. \tag{49}$$

Thus, if $r_y r_z$ is not in the interval $[0, 4]$, the solution is inadmissible. However, if $r_y r_z$ is in that interval the solution is not necessarily admissible.

Next we consider conditions that can ensure that the solution is admissible. Using (45), inequality (47) can be expressed as

$$w_{yt} + 2r_y \rho_t + r_y^2 w_{zt} > w_{yt}(1 - r_y r_z)^2. \tag{50}$$

Clearly this is satisfied if r_y, r_z and ρ_t have the same signs and $0 < r_y r_z < 1$, i.e., this would be one sufficient condition to ensure that the solution is admissible. An alternative sufficient condition can be obtained as follows. If $\tilde{\rho}_t$ is the residual correlation at time t,

$$w_{yt} + 2r_y \rho_t + r_y^2 w_{zt} > w_{yt}(1 - \tilde{\rho}_t^2). \tag{51}$$

Thus another sufficient condition for admissibility is

$$1 - \tilde{\rho}_t^2 > (1 - r_y r_z)^2 \tag{52}$$

or equivalently

$$r_y r_z (2 - r_y r_z) > \tilde{\rho}_t^2. \tag{53}$$

In conclusion, time-specific residual covariance can be added to the RCLPM as long as the reciprocal and the cross-lagged parameters are held time invariant. It is possible to further constrain the residual covariance or the residual correlation to be time invariant. Such a model will also be identified as it is nested within the above model.

1.5 The RCLPM without cross-lagged regressions

Consider the RCLPM without the cross-lagged regressions but with residual covariance. This model will be referred to as RLPM (the L refers to the lagged auto regression for each variable). The RLPM is given by the following equations

$$Y_t = a_{yt} + r_{yt} Z_t + b_{1t} Y_{t-1} + \varepsilon_{yt}$$

$$Z_t = a_{zt} + r_{zt} Y_t + b_{4t} Z_{t-1} + \varepsilon_{zt}$$

$$\varepsilon_{yt} \sim N(0, w_{yt})$$

$$\varepsilon_{zt} \sim N(0, w_{zt})$$

$$w_t = Cov(\varepsilon_{yt}, \varepsilon_{zt})$$

In matrix form the model is given by

$$\begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} + \begin{pmatrix} r_{yt} & 0 \\ 0 & r_{zt} \end{pmatrix} \begin{pmatrix} Y_t \\ Z_t \end{pmatrix} + \begin{pmatrix} b_{1t} & 0 \\ 0 & b_{4t} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} \tag{54}$$

where

$$Var \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} = \begin{pmatrix} w_{yt} & w_t \\ w_t & w_{zt} \end{pmatrix}.$$

This model is as usual converted to

$$\begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} + \tag{55}$$

$$\frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} b_{1t} & 0 \\ 0 & b_{4t} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \tag{56}$$

$$\frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix}. \tag{57}$$

We can see that the model has the same expression as the CLPM and it has the same number of parameters as the CLPM. In fact the two models are equivalent. The parameters of the CLPM can be obtained from the parameters of the RLPM using the following equations

$$\begin{pmatrix} \alpha_{yt} \\ \alpha_{zt} \end{pmatrix} = \frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} \tag{58}$$

$$\begin{pmatrix} \beta_{1t} & \beta_{2t} \\ \beta_{3t} & \beta_{4t} \end{pmatrix} = \frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} b_{1t} & 0 \\ 0 & b_{4t} \end{pmatrix} \tag{59}$$

$$\begin{pmatrix} v_{yt} & c_t \\ c_t & v_{zt} \end{pmatrix} = \frac{1}{(1 - r_{yt} r_{zt})^2} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} w_{yt} & w_t \\ w_t & v_{zt} \end{pmatrix} \begin{pmatrix} 1 & r_{zt} \\ r_{yt} & 1 \end{pmatrix} \tag{60}$$

These equations are easily reversible and we can obtain the parameters of the RLPM model from the parameters of the CLPM. The first step is to determine r_{yt} and r_{zt}. Equation (59) is equivalent to

$$\begin{pmatrix} 1 & -r_{yt} \\ -r_{zt} & 1 \end{pmatrix} \begin{pmatrix} \beta_{1t} & \beta_{2t} \\ \beta_{3t} & \beta_{4t} \end{pmatrix} = \begin{pmatrix} b_{1t} & 0 \\ 0 & b_{4t} \end{pmatrix}. \tag{61}$$

From here we obtain

$$r_{yt} = \frac{\beta_{2t}}{\beta_{4t}} \tag{62}$$

$$r_{zt} = \frac{\beta_{3t}}{\beta_{1t}}. \tag{63}$$

Equation (61) gives the expressions for b_{1t} and b_{4t}. Equations (58) and (60) are now also completely reversible

$$\begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} = \begin{pmatrix} 1 & -r_{yt} \\ -r_{zt} & 1 \end{pmatrix} \begin{pmatrix} \alpha_{yt} \\ \alpha_{zt} \end{pmatrix} \tag{64}$$

$$\begin{pmatrix} w_{yt} & w_t \\ w_t & v_{zt} \end{pmatrix} = \begin{pmatrix} 1 & -r_{yt} \\ -r_{zt} & 1 \end{pmatrix} \begin{pmatrix} v_{yt} & c_t \\ c_t & v_{zt} \end{pmatrix} \begin{pmatrix} 1 & -r_{zt} \\ -r_{yt} & 1 \end{pmatrix} \tag{65}$$

The equivalence between the two models is established as long as all denominators are not zero, i.e.,

$$\beta_{1t} \neq 0 \tag{66}$$

$$\beta_{4t} \neq 0 \tag{67}$$

$$\beta_{1t}\beta_{4t} \neq \beta_{2t}\beta_{3t} \tag{68}$$

or equivalently

$$b_{1t} \neq 0 \tag{69}$$

$$b_{4t} \neq 0 \tag{70}$$

$$r_{yt}r_{zt} \neq 1. \tag{71}$$

If these inequalities are satisfied not just by the point estimates but the entire posterior distributions, we can expect the RLPM to exhibit easy identifiability and approximately normal posterior distributions. Note here that the identification of the RLPM without cross-lags does not require equality constraints across-time, i.e., the reciprocal regression parameters can be time-specific.

Note that the absence of cross-lagged regressions in the RLPM avoids the dual solution problem discussed in Appendix Section 1.2 but it does not avoid the possible negative R^2 issue discussed in Appendix Section 1.3.