Dynamic Econometric Models

Use of lags is very common in Economics/Finance, particularly for the time-series data.

Reasons for Lags:

- Technological reasons
- Psychological reasons
- Institutional reasons

Type of Lag Models:

- **Distributed Lag Models** - $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots\ldots.. + \beta_k X_{t-k} + u_t$
- **Autoregressive Models** - $Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots\ldots.. + \beta_k Y_{t-k} + \gamma X_t + u_t$
- **Autoregressive Distributed Lag Models –**

  $Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots\ldots.. + \beta_k Y_{t-k} + \gamma_0 X_t + \gamma_1 X_{t-1} + \gamma 2 X_{t-2} + \ldots. + \gamma_m X_{t-m} + u_t$

Part-I

Distributed Lag Model:

$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots\ldots.. + \beta_k X_{t-k} + u_t$

Implications:

(a) ***Short-run/Impact Multiplier***: β_0

(b) ***Interim/Intermediate Multipliers***: $\beta_0 + \beta_1$, $\beta_0 + \beta_1 + \beta_2$, $\beta_0 + \beta_1 + \beta_2 + \beta_3$ ……….

(c) ***Long-run/Total Distributed Lag Multiplier***: $\beta = \beta_0 + \beta_1 + \beta_2 + \ldots.. + \beta_k = \sum_{j=0}^{k} B_j$

Important Issues:

(a) Selection of optimum lag length

(b) Introduction of more lags results loss of further degrees of freedom

(c) Possibility of severe multicollinearity

Estimation Procedure:

(a) <u>Adhoc Procedure</u>

- One needs to proceed sequentially, i.e., regressing Y_t on X_t, then Y_t on X_t and X_{t-1}, and so on.
- The sequential procedure stops when the regression coefficients of the lagged variables start becoming statistically NOT significant or sign of the coefficient of at least one of the lagged variables changes.
- However, there is no prior theory. Hence, it is a purely **adhoc** process.
- There will be loss of degrees of freedom with successive lags.
- There is also the possibility of the problem of severe multicollinearity.

(b) <u>Koyck Transformation Procedure:</u>

Model: $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + ... + u_t$ (1)

Assumption: $\beta_j = \beta_0 \lambda^j;\ 0 < \lambda < 1,\ j = 0,1,2,.....$ As $j \to \infty, \beta_j \to 0$ (2)

This means that β decreases geometrically as the lag length increases

Substituting (2) in (1), $Y_t = \alpha + \beta_0 X_t + \beta_0 \lambda X_{t-1} + \beta_0 \lambda^2 X_{t-2} + ... + u_t$ (3)

Lagging (3) by one period and multiplying both the sides by λ,

$\lambda Y_{t-1} = \alpha\lambda + \beta_0 \lambda X_{t-1} + \beta_0 \lambda^2 X_{t-2} + \beta_0 \lambda^3 X_{t-3} + ... + \lambda u_{t-1}$ (4)

Subtracting (4) from (3), $Y_t = \alpha^* + \beta_0 X_t + \lambda Y_{t-1} + v_t;\ v_t = u_t - \lambda u_{t-1};\ \alpha^* = \alpha(1-\lambda)$ (5)

- Koyck transformation reduces the number of coefficients to be estimated, but results in autoregressive model.
- Two potential problems - (i) Endogeneity (because of inclusion of Y_{t-1} as one of the independent variables), if v_t is autocorrelated, leading to biased and inconsistent OLS estimators; and (ii) Autocorrelation (if u_t is not autocorrelated) resulting in inefficient OLS estimators

- However, such transformation can solve the problem of autocorrelation if u_t is autocorrelated and λ approximates to the autocorrelation coefficients.

- Application of **instrumental variable(s) for estimation of the model–** Using a proxy variable for the endogenous independent variable (Y_{t-1} here), which is strongly correlated with this endogenous variable, but not so with the random disturbance term (v_t).

(c) <u>Almon Transformation Procedure:</u>

It is assumed that β_j is a suitable degree polynomial function of j, i.e.,

$$\beta_j = a_0 + a_1 j + a_2 j^2;\ \beta_j = a_0 + a_1 j + a_2 j^2 + a_3 j^3;\ \beta_j = a_0 + a_1 j + a_2 j^2 + a_3 j^3 + ...+ a_m j^m$$

It is based on the **Weierstrass' Theorem** (on a finite closed interval, any continuous function may be approximated uniformly by a polynomial of a suitable degree).

An Example:

Model: $$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + ...+ \beta_k X_{t-k} + u_t = \alpha + \sum_{j=0}^{k} \beta_j X_{t-j} + u_t \qquad (1)$$

Assumption: $\beta_j = a_0 + a_1 j + a_2 j^2$ (i.e., β_j is a quadratic function of j) (2)

Substituting (2) in (1), $Y_t = \alpha + \sum_{j=0}^{k} (a_o + a_1 j + a_2 j^2) X_{t-j} + u_t$

Or, $$Y_t = \alpha + a_0 \sum_{j=0}^{k} X_{t-j} + a_1 \sum_{j=0}^{k} j.X_{t-j} + a_2 \sum_{j=0}^{k} j^2.X_{t-j} + u_t = \alpha + a_0 Z_{0t} + a_1 Z_{1t} + a_2 Z_{2t} + u_t \qquad (3)$$

Here, $$Z_{0t} = \sum_{j=0}^{k} X_{t-j},\ Z_{1t} = \sum_{j=0}^{k} j.X_{t-j},\ Z_{2t} = \sum_{j=0}^{k} j^2.X_{t-j} \qquad (4)$$

- The original coefficients may be derived in the following way:

$$\hat{\beta}_0 = \hat{a}_0;\ \hat{\beta}_1 = \hat{a}_0 + \hat{a}_1 + \hat{a}_2;\ \hat{\beta}_2 = \hat{a}_0 + 2\hat{a}_1 + 4\hat{a}_2;\ \hat{\beta}_3 = \hat{a}_0 + 3\hat{a}_1 + 9\hat{a}_2;\ ...;\ \hat{\beta}_k = \hat{a}_0 + k\hat{a}_1 + k^2\hat{a}_2$$

- The model can be estimated by applying the OLS method.

- However, it requires specification of lag length and the degree of polynomial.
  - ✓ Selection of the lag length can be based on the underlying theory and/or different information criteria (e.g., Akaike Information Criterion - AIC,

Schwartz Bayesian Information Criterion-SBIC, etc.). One can also select the lag length that has the highest R^2.

- ✓ The degree of polynomial can be selected sequentially starting with the highest degree polynomial possible and going backward until the null hypothesis of one of the coefficients is rejected. However, one can also select the degree of polynomial on the basis of the underlying theory, if any.

- There may be multicollinearity problem among Z, though it may not be serious for the original coefficients.

Part-II

Models with unobservable dependent and/or independent variable(s)

(1) Application of Adaptive Expectations Hypothesis

Model: $Y_t = \alpha + \beta X^*_{t+1} + u_t$ (1)

The Adaptive Expectations Hypothesis

It is assumed that expectations are revised based on most recent error, i.e.,

$$(X^*_{t+1} - X^*_t) = \theta(X_t - X^*_t)\,;\ 0 < \theta < 1 \qquad (2)$$

Here, $(X^*_{t+1} - X^*_t)$ = Revision in expectation, $(X_t - X^*_t)$ = Last period's error; and θ = Expectation adjustment coefficient

It is also known as the **progressive expectation** or **error learning hypothesis** propounded by Cagan (1956) and Friedman (1957)

The hypothesis can also be written as,

$$X^*_{t+1} = X^*_t + \theta(X_t - X^*_t) = \theta X_t + (1-\theta)X^*_t \qquad (3)$$

Here, expected X for period t+1 is a weighted average of actual (observed) X and expected X at period t with θ and (1-θ) respectively being the weights.

Now, (3) can be rewritten as $X^*_{t+1} - (1-\theta)X^*_t = \theta X_t$ (4)

Lagging (1) by 1 period and multiplying both the sides by $(1-\theta)$,

$$(1-\theta)Y_{t-1} = \alpha(1-\theta) + \beta(1-\theta)X_t^* + (1-\theta)u_{t-1} \quad (5)$$

Subtracting (5) from (1), $Y_t - (1-\theta)Y_{t-1} = \alpha\theta + \beta[X_{t+1}^* - (1-\theta)X_t^*] + [u_t - (1-\theta)u_{t-1}]$ (6)

Substituting (4) in (6),

$$Y_t = \alpha\theta + (1-\theta)Y_{t-1} + \beta\theta X_t + v_t \text{ with } v_t = u_t - (1-\theta)u_{t-1} \quad (7)$$

Or, $Y_t = \alpha^* + \theta^* Y_{t-1} + \beta^* X_t + v_t$ with $\alpha^* = \alpha\theta$, $\theta^* = (1-\theta)$ and $\beta^* = \beta\theta$ (8)

- This is similar to the equation obtained by using Koyck transformation procedure.
- Three coefficients are to be estimated from the derived equation – α^*, β^* and θ^*. The coefficients of the original equation can be derived from these estimated coefficients.
- The derived model is autoregressive in nature.
- Two potential problems - (i) Endogeneity (because of inclusion of Y_{t-1} as one of the independent variables), if v_t is autocorrelated, leading to biased and inconsistent OLS estimators; and (ii) Autocorrelation (if u_t is not autocorrelated) resulting in inefficient OLS estimators
- However, such transformation can solve the problem of autocorrelation if u_t is autocorrelated and (1-θ) approximates to the autocorrelation coefficients.
- Application of **instrumental variable(s) for estimation of the model –** Using a proxy variable for the endogenous independent variable (Y_{t-1} here), which is strongly correlated with this endogenous variable, but not with the random disturbance term.

(2) **Application of Partial Adjustment Hypothesis**

Model: $Y_t^* = \alpha + \beta X_t + u_t$ (1)

The Partial Adjustment Hypothesis

Assumes that the variables are adjusted only partially toward their desired levels, i.e.,

$$(Y_t - Y_{t-1}) = \lambda(Y_t^* - Y_{t-1});\ 0 < \lambda < 1 \quad (2)$$

Here, $(Y_t - Y_{t-1})$ = Actual change, $(Y_t^* - Y_{t-1})$ = Desired change, λ = Coefficient of adjustment

Now (2) can be rewritten as $Y_t = \lambda Y_t^* + (1-\lambda)Y_{t-1}$ (3)

This means that observed Y at time t is a weighted average of desired Y at that time and actual Y at the previous period with λ and (1-λ) respectively being the weights

It is also known as the **Stock Adjustment Hypothesis** as propounded by Nerlove (1958).

When $Y_t = X_{t+1}^*$ and $Y_t^* = X_t$, the Partial Adjustment Hypothesis is similar to the Adaptive Expectation Hypothesis.

Substituting (1) in (3), $(Y_t - Y_{t-1}) = \lambda(\alpha + \beta X_t + u_t - Y_{t-1})$

Or, $Y_t = \lambda\alpha + (1-\lambda)Y_{t-1} + \lambda\beta X_t + \lambda u_t$

Or, $Y_t = \alpha^* + \lambda^* Y_{t-1} + \beta^* X_t + v_t$ with $\alpha^* = \alpha\lambda$; $\lambda^* = (1-\lambda)$; $\beta^* = \beta\lambda$ and $v_t = \lambda u_t$ (4)

- Like the model with Koyck transformation or the model with adaptive expectation hypothesis, this derived model is also autoregressive.
- Three coefficients are to be estimated from the derived equation – α^*, β^* and λ^*. The coefficients of the original equation can be derived from these estimated coefficients.
- There is potential problem of **endogeneity** (because of inclusion of Y_{t-1} as one of the independent variables), if v_t is autocorrelated, leading to biased and inconsistent OLS estimators.
- However, there will be no problem of autocorrelation in the derived model, if u_t is not autocorrelated (as $v_t = \lambda u_t$)
- On contrary, if u_t is autocorrelated, the transformed model will also have the problem of autocorrelation resulting in inefficient OLS estimators.
- Application of **instrumental variable(s)** for estimation of the model **–** Using a proxy variable for the endogenous independent variable (Y_{t-1} here), which is strongly correlated with this endogenous variable, but not with the random disturbance term.

(3) Application of Adaptive Expectations Hypothesis and Partial Adjustment Hypothesis

Model: $Y_t^* = \alpha + \beta X_{t+1}^* + u_t$ (1)

Adaptive expectation hypothesis: $(X^*_{t+1} - X^*_t) = \theta(X_t - X^*_t)$ (2)

Partial Adjustment Hypothesis: $(Y_t - Y_{t-1}) = \lambda(Y^*_t - Y_{t-1})$ (3)

Following the earlier processes,

$$Y_t = \alpha\theta\lambda + \beta\theta\lambda X_t + [(1-\theta) + (1-\lambda)]Y_{t-1} - (1-\theta)(1-\lambda)Y_{t-2} + \lambda[u_t - (1-\theta)u_{t-1}] \quad (4)$$

$$Y_t = \gamma_0 + \gamma_1 X_t + \gamma_2 Y_{t-1} + \gamma_3 Y_{t-2} + v_t \quad (5)$$

Here, $\gamma_0 = \alpha\theta\lambda$, $\gamma_1 = \beta\theta\lambda$, $\gamma_2 = [(1-\theta) + (1-\lambda)]$, $\gamma_3 = -(1-\theta)(1-\lambda)$, and $v_t = \lambda[u_t - (1-\theta)u_{t-1}]$

- Four coefficients are to be estimated from the derived equation – γ_0, γ_1, γ_2, and γ_3. The coefficients of the original equation can be derived from these estimated coefficients.
- Like the models with Koyck transformation, the model with the adaptive expectation hypothesis or the model with the partial adjustment hypothesis, this derived model is also autoregressive, but of second order.
- There is potential problem of **endogeneity** (because of inclusion of Y_{t-1} and Y_{t-2} as the independent variables), if v_t is autocorrelated, leading to biased and inconsistent OLS estimators.
- There will be problem of autocorrelation in the derived model (if u_t is not autocorrelated) resulting in inefficient OLS estimators. However, such transformation can solve the problem of autocorrelation if u_t is autocorrelated and (1-θ) approximates to the autocorrelation coefficients.
- Application of **instrumental variables** for estimation of the model **–** Using a proxy variable for the endogenous independent variable (Y_{t-1} and Y_{t-2} here), which are strongly correlated with these endogenous variable, but not so with the random disturbance term.

<u>Problem of Lack of Uniqueness in Solutions in ARDL Models: Example</u>

Model: $Y^*_t = \alpha + \beta_1 X^*_t + \beta_2 Z_t + u_t$ (1)

Adaptive expectation hypothesis: $(X^*_t - X^*_{t-1}) = \theta(X_{t-1} - X^*_{t-1})$ or $[X^*_t - (1-\theta)X^*_{t-1}] = \theta X_{t-1}$ (2)

Partial Adjustment Hypothesis: $(Y_t - Y_{t-1}) = \lambda(Y^*_t - Y_{t-1})$ or $Y_t = \lambda Y^*_t + (1-\lambda)Y_{t-1}$ (3)

Substituting (1) in (3), $Y_t = \lambda(\alpha + \beta_1 X_t^* + \beta_2 Z_t + u_t) + (1-\lambda)Y_{t-1}$ (4)

Or, $Y_t = \lambda\alpha + \beta_1\lambda X_t^* + \beta_2\lambda Z_t + \lambda u_t + (1-\lambda)Y_{t-1}$ (5)

Lagging (5) by one period and multiplying both the sides by $(1-\theta)$,

$$(1-\theta)Y_{t-1} = \lambda\alpha(1-\theta) + \beta_1\lambda(1-\theta)X_{t-1}^* + \beta_2\lambda(1-\theta)Z_{t-1} + \lambda(1-\theta)u_{t-1} + (1-\theta)(1-\lambda)Y_{t-2} \quad (6)$$

Subtracting (6) from (5),

$$\begin{aligned} Y_t &= \lambda\alpha\theta + \beta_1\lambda[X_t^* - (1-\theta)X_{t-1}^*] + \beta_2\lambda Z_t + \beta_2\lambda(1-\theta)Z_{t-1} + [(1-\lambda) + (1-\theta)]Y_{t-1} \\ &- (1-\theta)(1-\lambda)Y_{t-2} + \lambda[u_t - (1-\theta)u_{t-1}] \end{aligned} \quad (7)$$

Or, $$\begin{aligned} Y_t &= \lambda\alpha\theta + \beta_1\lambda\theta X_{t-1} + \beta_2\lambda Z_t - \beta_2\lambda(1-\theta)Z_{t-1} + [(1-\lambda) + (1-\theta)]Y_{t-1} \\ &- (1-\theta)(1-\lambda)Y_{t-2} + v_t \end{aligned} \quad (8)$$

Or, $Y_t = \alpha_1 + \alpha_2 X_{t-1} + \alpha_3 Z_t + \alpha_4 Z_{t-1} + \alpha_5 Y_{t-1} + \alpha_6 Y_{t-2} + v_t$ (9)

Here, $\alpha_1 = \lambda\alpha\theta$, $\alpha_2 = \beta_1\lambda\theta$, $\alpha_3 = \beta_2\lambda$, $\alpha_4 = -\beta_2\lambda(1-\theta)$, $\alpha_5 = [(1-\lambda) + (1-\theta)]$ $\alpha_6 = -(1-\lambda)(1-\theta)$, and $v_t = \lambda[u_t - (1-\theta)u_{t-1}]$

Thus, we have **six equations (α_1, α_2, α_3, α_4, α_5, and α_6)** to solve for **five unknowns (α, β_1, β_2, λ and θ)** – There is problem of lack of uniqueness in the solutions.

Example: $\theta = 1 + \frac{\alpha_4}{\alpha_3}$, but we get two estimates of λ - (1) $\lambda = 2 - \alpha_5 - \theta$, and (2) $\lambda = 1 + \frac{\alpha_6}{1-\theta}$ for the derived value of θ.

How to overcome the problem of lack of uniqueness?

We can rewrite (8) as,

$$Y_t - (1-\theta)Y_{t-1} = \lambda\alpha\theta + \beta_1\lambda\theta X_{t-1} + \beta_2\lambda[Z_t - (1-\theta)Z_{t-1}] + (1-\lambda)[Y_{t-1} - (1-\theta)Y_{t-2}] + v_t \quad (10)$$

Or, $\tilde{Y}_t = \lambda\alpha\theta + \beta_1\lambda\theta X_{t-1} + \beta_2\lambda\tilde{Z}_t + (1-\lambda)\tilde{Y}_{t-1} + v_t$ (11)

Or, $\tilde{Y}_t = \delta_0 + \delta_1 X_{t-1} + \delta_2\tilde{Z}_t + \delta_3\tilde{Y}_{t-1} + v_t$ (12)

$\tilde{Y}_t = Y_t - (1-\theta)Y_{t-1}$; $\tilde{Z}_t = Z_t - (1-\theta)Z_{t-1}$, and $\tilde{Y}_{t-1} = Y_t - (1-\theta)Y_{t-2}$

$\delta_0 = \lambda\alpha\theta$; $\delta_1 = \beta_1\lambda\theta$; $\delta_2 = \beta_2\lambda$; and $\delta_3 = 1-\lambda$

Here, we have **four equations (δ_0, δ_1, δ_2 and δ_3)** which can be used to solve **four unknown coefficients (α, β_1, β_2 and λ)** – leading to uniqueness in solutions (given the derived θ in Step 1)

Part-III

Granger Causality Test

Direction of causal relationships is very important in Econometric modelling. Given the past values of variable *Y*, if the past values of variable *X* are useful for predicting *Y*, variable *X* is said to Granger cause variable *Y*.

Basic Equations:

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{j=1}^{q} \gamma_j X_{t-j} + u_t \text{ and } X_t = \gamma + \sum_{i=1}^{r} \delta_i X_{t-i} + \sum_{i=1}^{s} \theta_i Y_{t-j} + v_t$$

Four Possible Cases:

1. ***Unidirectional causality from X to Y***
   - If the coefficients of lagged *X* in the first equation are statistically significant, whereas those of lagged Y are not as a group

2. **Unidirectional causality *from Y to X***
   - If the coefficients of lagged *Y* in the second equation are statistically significant, whereas those of lagged *X* are not as a group

3. ***Feedback or bilateral causality***
   - If the coefficients of both lagged X and lagged Y are statistically significant in both the equations

4. ***Independence***
   - If the coefficients of both lagged X and lagged Y are not statistically significant in both the equations

Basic Approach Followed:

(a) Testing if X Granger Causes Y

- Regressing current Y on lagged Y, but without inclusion of X – This gives the **Restricted Residual Sum of Squares (RRSS)**

- Regressing current Y on both lagged Y and lagged X – This gives the **Unrestricted Residual Sum of Squares (URSS)**
- Null Hypothesis: Lagged X terms do not influence Y
- Carrying out Restricted F test
- Rejection of the Null Hypothesis implies that X Granger causes Y

(b) Testing if Y Granger Causes X

- Regressing current X on lagged X, but without inclusion of Y – This gives the **Restricted Residual Sum of Squares (RRSS)**
- Regressing current X on both lagged X and lagged Y – This gives the **Unrestricted Residual Sum of Squares (URSS)**
- Null Hypothesis: Lagged Y terms do not influence X
- Carrying out Restricted F test
- Rejection of the Null Hypothesis implies that Y Granger causes X

(c) Testing if there is Feedback or Bidirectional Causality

- Rejection of the Null Hypothesis in both (a) and (b) implies that there is feedback or bidirectional causality.

(d) Testing if X and Y are Independent

- Non-Rejection of the Null Hypothesis in both (a) and (b) implies that X and Y independent.

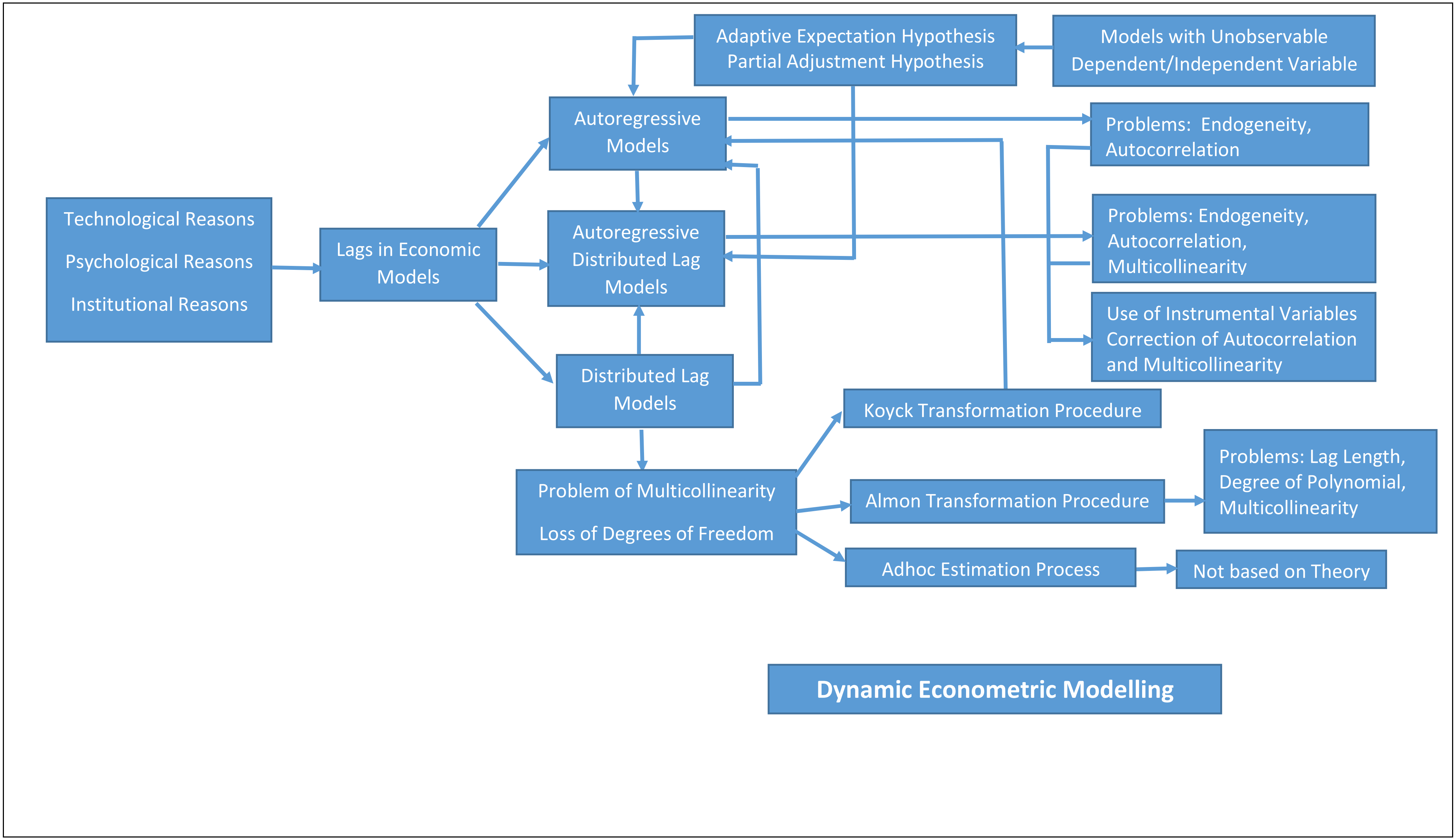
Some Cautions:

- It is necessary to ensure the ***stationary*** nature of the variables.
- Selection of the lag length may influence the direction of causality.
- The error terms in the two equations must not be autocorrelated.

Granger causality/non-causality vs. Exogeneity

Granger causality is neither necessary nor sufficient to establish exogeneity

- ***Weak Exogeneity of X*** – If current Y does not explain current X – Estimation and Inference can be done with weak exogeneity.
- ***Strong Exogeneity of X*** – If current as well as past Y do not explain X – Strong exogeneity is Important for prediction/forecasting.
- ***Super Exogeneity of X –*** No change in the parameters in the regression of Y on X even if X changes (i.e., no structural break or instability in the parameters)- This is very important for policy formulation.


Adaptive Expectation Hypothesis
Partial Adjustment Hypothesis
Models with Unobservable Dependent/Independent Variable
Autoregressive Models
Problems: Endogeneity, Autocorrelation
Technological Reasons
Psychological Reasons
Institutional Reasons
Lags in Economic Models
Autoregressive Distributed Lag Models
Problems: Endogeneity, Autocorrelation, Multicollinearity
Use of Instrumental Variables
Correction of Autocorrelation and Multicollinearity
Distributed Lag Models
Koyck Transformation Procedure
Problem of Multicollinearity
Loss of Degrees of Freedom
Almon Transformation Procedure
Problems: Lag Length, Degree of Polynomial, Multicollinearity
Adhoc Estimation Process
Not based on Theory
Dynamic Econometric Modelling