Econ 582
Dynamic Regression Models

Eric Zivot

April 29, 2013

Distributed Lag (DL) Models

Consider the stylized regression model with a single lagged variable

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$
$$x_t \text{ is covariance stationary and ergodic}$$
$$x_t \text{ is strictly exogenous, } E[\varepsilon_t|x_s] = 0 \text{ for all } t, s$$

Remarks:

1. Covariance stationarity of x_t implies

$$x_t = \mu_x + \psi_x(L)v_t,\ v_t \sim WN(0, \sigma_v^2)$$

2. $E[\varepsilon_t|x_s] = 0$ for all t, s implies that $E[\varepsilon_t|v_s] = 0$ for all t, s

Example: Dynamic demand function

$$\begin{aligned} y_t &= \text{log quantity} \\ x_t &= \text{log price} \end{aligned}$$

Dynamic multipliers (impulse responses)

$$\begin{aligned} \frac{\partial y_t}{\partial x_t} &= \beta_0 = \text{contemporaneous price elasticity} \\ \frac{\partial y_t}{\partial x_{t-1}} &= \frac{\partial y_{t+1}}{\partial x_t} = \beta_1 = \text{lag 1 price elasticity} \\ \frac{\partial y_t}{\partial x_{t-j}} &= \frac{\partial y_{t+j}}{\partial x_t} = 0 = \text{lag j price elasticity } (j > 1) \end{aligned}$$

Cumulative or long-run price elasticity

$$\lim_{j \to \infty} \sum_{j=0}^{\infty} \frac{\partial y_{t+j}}{\partial x_t} = \beta_0 + \beta_1$$

Alternative Interpretation of Long-run Price Elasticity

Suppose x_t is at its steady-state value (unconditional mean): $x_t = x_{t-1} = \mu_x$ and $\varepsilon_t = 0$. Define the long-run value for y_t as its unconditional mean

$$\begin{aligned} E[y_t] &= \alpha + \beta_0 E[x_t] + \beta_1 E[x_{t-1}] + E[\varepsilon_t] \\ &\Rightarrow \mu_y = \alpha + \beta_0 \mu_x + \beta_1 \mu_x = \alpha + (\beta_0 + \beta_1)\mu_x \end{aligned}$$

Note that

$$\frac{\partial \mu_y}{\partial \mu_x} = \beta_0 + \beta_1$$

Lag Operator Notation

$$\begin{aligned} y_t &= \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \\ &= \alpha + (\beta_0 + \beta_1 L) x_t = \varepsilon_t \\ &= \alpha + \beta(L) x_t + \varepsilon_t, \\ \beta(L) &= \beta_0 + \beta_1 L \end{aligned}$$

Note

$$\frac{\partial \mu_y}{\partial \mu_x} = \beta_0 + \beta_1 = \beta(1)$$

General DL(p) Model

$$
\begin{aligned}
y_t &= \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p} + \varepsilon_t, \\
& \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2) \\
E[\varepsilon_t | x_s] &= 0 \text{ for all } t, s
\end{aligned}
$$

In lag notation

$$
y_t = \alpha + \beta(L) x_t + \varepsilon_t, \ \beta(L) = \beta_0 + \beta_1 L + \cdots + \beta_p L^p
$$

Dynamic multipliers

$$
\begin{aligned}
\frac{\partial y_t}{\partial x_{t-j}} &= \frac{\partial y_{t+j}}{\partial x_t} = \beta_j \text{ for } j < p \\
\frac{\partial \mu_y}{\partial \mu_x} &= \beta_0 + \beta_1 + \cdots + \beta_p = \beta(1)
\end{aligned}
$$

Estimation of DL(p) Models

- OLS estimation produces unbiased, consistent and asymptotically normally distributed estimates provides x_t is covariance stationary and ergodic, $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ and x_t is strictly exogenous.

- If $\varepsilon_t \sim$iid $N(0, \sigma_\varepsilon^2)$ then

$$y_t | x_t, \ldots, x_{t-p} \sim N(\alpha + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p}, \sigma_\varepsilon^2)$$

  and OLS is equivalent to conditional MLE

- Main specification issue is how to choose the lag length. Note: x_t, x_{t-1}, $\ldots$, x_{t-p} may be highly correlated.

Lag Length Specification

General-to-Specific Strategy (backwards selection)

1. Specify p_{max} and estimate DL model by OLS

$$y_t = \hat{\alpha} + \hat{\beta}_0 x_t + \hat{\beta}_1 x_{t-1} + \cdots + \hat{\beta}_p x_{t-p} + \hat{\varepsilon}_t$$

2. Test significance of x_{t-p} using t-tests and test joint significance of $x_{t-k}, \ldots, x_{t-p}$ using F-tests

3. Eliminate insignificant regressors and repeat testing until all regressors are significant

Model Selection Criteria

1. Specify p_{max} and estimate all DL(p) models for $p \leq p_{\mathsf{max}}$ by OLS using common sample $t = p_{\mathsf{max}} + 1, \ldots, T$. Define $N = T - (p_{\mathsf{max}} + 1)$.

2. Compute AIC(p) and/or BIC(p) for each model where

$$
\begin{aligned}
AIC(p) &= \ln \hat{\sigma}_{\varepsilon}^{2} + N^{-1} 2p \\
BIC &= \ln \hat{\sigma}_{\varepsilon}^{2} + \frac{\ln N}{N} p
\end{aligned}
$$

3. Best model is one with smallest AIC or BIC

Dealing with High Correlation among Regressors

Problem: $x_t, x_{t-1}, \ldots, x_{t-p}$ may be highly correlated (e.g, x_t follows AR(1) with ϕ close to 1)

Implication: In estimated regression

$$y_t = \hat{\alpha} + \hat{\beta}_0 x_t + \hat{\beta}_1 x_{t-1} + \cdots + \hat{\beta}_p x_{t-p} + \hat{\varepsilon}_t,$$

individual $\hat{\beta}_i$ may not be significant but jointly $\hat{\beta}_1, \ldots, \hat{\beta}_p$ are significant. This can screw up model selection.

Remark: Reparameterization can sometimes help

Example: DL(1)

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

Problem: x_t and x_{t-1} are highly correlated because x_t is highly autocorrelated.

Solution: Reparametrize model (add and subtract $\beta_0 x_{t-1}$ from RHS)

$$\begin{aligned} y_t &= \alpha + \beta_0 x_t - \beta_0 x_{t-1} + \beta_1 x_{t-1} + \beta_0 x_{t-1} + \varepsilon_t \\ &= \alpha + \beta_0 \Delta x_t + (\beta_0 + \beta_1) x_{t-1} + \varepsilon_t \\ &= \alpha + \beta_0 \Delta x_t + \beta(1) x_{t-1} + \varepsilon_t \end{aligned}$$

OLS on original and reparametrized model will give the same fit. However, in reparameterized model Δx_t and x_{t-1} will have low correlation

Also: In reparameterized model short-run response, β_0, and long-run response, $\beta(1)$, are directly estimated.

The Geometric Lag Model

Idea: Impose smoothness restrictions in lag structure to reduce number of estimated parameters

$$\begin{aligned} \text{DL}(\infty) \quad &: \quad y_t = \alpha + \sum_{k=0}^{\infty} \beta_k x_{t-k} + \varepsilon_t \\ \beta_k \quad &= \quad \beta(1-\lambda)\lambda^k, \ 0 \leq \lambda < 1 \end{aligned}$$

Here, λ measures persistence of lags (decay rate of impulse response function)

- $\lambda \approx 0 \Rightarrow$ low persistence (fast decay of impulse responses)
- $\lambda \approx 1 \Rightarrow$ high persistence (slow decay of impulse responses)

Substituting $\beta_k = \beta(1-\lambda)\lambda^k$ into DL(∞) gives

$$\begin{aligned} y_t &= \alpha + \sum_{k=0}^{\infty} \beta_k x_{t-k} + \varepsilon_t \\ &= \alpha + \beta \sum_{k=0}^{\infty} (1-\lambda)\lambda^k x_{t-k} + \varepsilon_t \\ &= \alpha + \beta \cdot B(L) x_t + \varepsilon_t \\ B(L) &= \sum_{k=0}^{\infty} b_k L^k = \sum_{k=0}^{\infty} (1-\lambda)\lambda^k L^k = (1-\lambda)(1-\lambda L)^{-1} \end{aligned}$$

Note

$$B(1) = \sum_{k=0}^{\infty} b_k = (1-\lambda) \sum_{k=0}^{\infty} \lambda^k = 1$$

Impulse Responses (Dynamic Multipliers)

$$
\begin{aligned}
y_t &= \alpha + \beta \sum_{k=0}^{\infty} (1-\lambda)\lambda^k x_{t-k} + \varepsilon_t \\
&= \alpha + \beta(1-\lambda)x_t + \cdots + \beta(1-\lambda)\lambda^k x_{t-k} + \cdots + \varepsilon_t
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial y_t}{\partial x_t} &= \beta(1-\lambda) < \beta \\
\frac{\partial y_t}{\partial x_{t-k}} &= \frac{\partial y_{t+k}}{\partial x_t} = \beta(1-\lambda)\lambda^k \\
\sum_{k=0}^{\infty} \frac{\partial y_{t+k}}{\partial x_t} &= \beta(1-\lambda) \sum_{k=0}^{\infty} \lambda^k = \beta
\end{aligned}
$$

Estimation of Geometric Lag Model

$$\begin{aligned} y_t &= \alpha + \beta \cdot B(L)x_t + \varepsilon_t \\ &= \alpha + \beta \cdot (1-\lambda)(1-\lambda L)^{-1}x_t + \varepsilon_t \end{aligned}$$

Multiplying both sides by $(1-\lambda L)$ gives the so-called Koyck transformed model

$$\begin{aligned} (1-\lambda L)y_t &= (1-\lambda L)\alpha + \beta \cdot (1-\lambda)x_t + (1-\lambda L)\varepsilon_t \Rightarrow \\ y_t &= (1-\lambda)\alpha + \lambda y_{t-1} + \beta \cdot (1-\lambda)x_t + \varepsilon_t - \lambda\varepsilon_{t-1} \\ &= \delta_0 + \lambda y_{t-1} + \delta_1 x_t + u_t \end{aligned}$$

where

$$\begin{aligned} \delta_0 &= (1-\lambda)\alpha,\ \delta_1 = \beta \cdot (1-\lambda) \\ u_t &= \varepsilon_t - \lambda\varepsilon_{t-1} = MA(1) \text{ process} \end{aligned}$$

Problem: In

$$\begin{aligned} y_t &= \delta_0 + \lambda y_{t-1} + \delta_1 x_t + u_t \\ u_t &= \varepsilon_t - \lambda \varepsilon_{t-1} \end{aligned}$$

we have

$$cov(y_{t-1}, u_t) \neq 0$$

Hence, OLS estimation gives biased and inconsistent estimates!

- Consistent estimation can be done by Instrumental Variables (IV) using y_{t-2} as an instrument for y_{t-1}.

Economic Models that Generate Geometric Lag Behavior

Geometric lag behavior can result from a variety of sources. Two common sources are

1. Adaptive expectations

2. Partial adjustment

Adaptive Expectations

Suppose the theoretical model is of the form

$$\begin{aligned} y_t &= \alpha + \beta x_t^e + \varepsilon_t \\ x_t^e &= \text{unobserved expectation variable} \\ &= E[x_{t+1}|I_t] \end{aligned}$$

Adaptive expectations:

$$\begin{aligned} \Delta x_t^e &= (1-\lambda)(x_t - x_{t-1}^e),\ 0 \leq \lambda < 1 \\ \Delta x_t^e &= \text{revision in expectations} \\ (1-\lambda) &= \text{speed of adjustment} \\ (x_t - x_{t-1}^e) &= x_t - E[x_t|I_{t-1}] = \text{ forecast error} \end{aligned}$$

Alternatively

$$\begin{aligned} x_t^e &= x_{t-1}^e + (1-\lambda)x_t - (1-\lambda)x_{t-1}^e \\ &= \lambda x_{t-1}^e + (1-\lambda)x_t \end{aligned}$$

Speed of adjustment (expectations revision)

- Instantaneous revision: $\lambda = 0 \Rightarrow \Delta x_t^e = (x_t - x_{t-1}^e) \Rightarrow x_t^e = x_t$
- No revision: $\lambda = 1 \Rightarrow \Delta x_t^e = 0 \Rightarrow x_t^e = x_{t-1}^e$
- Fast revision: λ close to zero
- Slow revision: λ close to one

Solving the Adaptive Expectations Model

$$\begin{aligned} x_t^e &= \lambda x_{t-1}^e + (1-\lambda)x_t \Rightarrow \\ (1-\lambda L)x_t^e &= (1-\lambda)x_t \Rightarrow \\ x_t^e &= (1-\lambda L)^{-1}(1-\lambda)x_t \\ &= (1-\lambda)\sum_{k=0}^{\infty}\lambda^k x_{t-k} \\ &= \sum_{k=0}^{\infty} b_k x_{t-k},\ b_k = (1-\lambda)\lambda^k \end{aligned}$$

Hence, x_t^e is an exponentially weighted average of current and past x values

Result: Economic model with adaptive expecations is a geometric DL model

$$\begin{aligned} y_t &= \alpha + \beta x_t^e + \varepsilon_t \\ x_t^e &= \lambda x_{t-1}^e + (1-\lambda) x_t \\ &= (1-\lambda) \sum_{k=0}^{\infty} \lambda^k x_{t-k} \end{aligned}$$

which implies

$$\begin{aligned} y_t &= \alpha + \beta(1-\lambda) \sum_{k=0}^{\infty} \lambda^k x_{t-k} + \varepsilon_t \\ &= \alpha + \beta B(L) x_t + \varepsilon_t \\ B(L) &= (1-\lambda)(1-\lambda L)^{-1} \end{aligned}$$

Partial Adjustment Model

Idea: In the absence of adjustment costs, the desired or equilibrium level of y_t is described by

$$y_t^* = \alpha + \beta x_t$$

However, if $y_t \neq y_t^*$ there is some cost required to adjust to the desired level and so each period there is only a partial adjustment

$$\begin{aligned} \Delta y_t &= (1-\lambda)(y_t^* - y_{t-1}) + \varepsilon_t,\ 0 < \lambda < 1, \varepsilon_t \sim WN(0, \sigma^2) \\ E[\Delta y_t] &= (1-\lambda)(y_t^* - y_{t-1}) \end{aligned}$$

Note:

- $y_{t-1} = y_t^* \Rightarrow E[\Delta y_t] = 0$

- $y_{t-1} \neq y_t^* \Rightarrow E[\Delta y_t] \neq 0$ and λ measures speed of adjustment

  - $\lambda \approx 0 \Rightarrow$ fast adjustment

  - $\lambda \approx 1 \Rightarrow$ slow adjustment

Result: Adjustment cost model is a linear regression with a lagged dependent variable

$$
\begin{aligned}
y_t^* &= \alpha + \beta x_t \\
\Delta y_t &= (1-\lambda)(y_t^* - y_{t-1}) + \varepsilon_t,\ 0 < \lambda < 1, \varepsilon_t \sim WN(0, \sigma^2)
\end{aligned}
$$

Express second equation in terms of y_t

$$
\begin{aligned}
y_t &= y_{t-1} - (1-\lambda)y_{t-1} + (1-\lambda)y_t^* + \varepsilon_t \\
&= \lambda y_{t-1} + (1-\lambda)y_t^* + \varepsilon_t
\end{aligned}
$$

Substitute $y_t^* = \alpha + \beta x_t$ to give

$$
\begin{aligned}
y_t &= \lambda y_{t-1} + (1-\lambda)(\alpha + \beta x_t) + \varepsilon_t \\
&= \alpha(1-\lambda) + \lambda y_{t-1} + \beta(1-\lambda)x_t + \varepsilon_t \\
&= \gamma + \lambda y_{t-1} + \pi x_t + \varepsilon_t
\end{aligned}
$$

Interpretation of Coefficients in Partial Adjustment Model

$$\begin{aligned} y_t^* &= \alpha + \beta x_t \\ y_t &= \alpha(1-\lambda) + \lambda y_{t-1} + \beta(1-\lambda)x_t + \varepsilon_t \\ &= \gamma + \lambda y_{t-1} + \pi x_t + \varepsilon_t \\ \gamma &= \alpha(1-\lambda),\ \pi = \beta(1-\lambda) \end{aligned}$$

Dynamic multipliers

$$\begin{aligned} \frac{\partial y_t}{\partial x_t} &= \pi = \beta(1-\lambda) \leq \beta \text{ bc } 0 \leq \lambda < 1 \\ &= \text{ short-run impact} \\ \frac{\partial y_t^*}{\partial x_t} &= \beta = \text{ long-run impact} \end{aligned}$$

Interpretation of long-run impact

$$
\begin{aligned}
y_t &= \gamma + \lambda y_{t-1} + \pi x_t + \varepsilon_t \Rightarrow \\
(1 - \lambda L) y_t &= \gamma + \pi x_t + \varepsilon_t \Rightarrow \\
y_t &= (1 - \lambda L)^{-1} \gamma + (1 - \lambda L)^{-1} \pi x_t + (1 - \lambda L)^{-1} \varepsilon_t \\
&= \frac{\gamma}{(1 - \lambda)} + \pi \sum_{k=0}^{\infty} \lambda^k x_{t-k} + \sum_{k=0}^{\infty} \lambda^k \varepsilon_{t-k} \\
&= \alpha + \pi \left(x_t + \lambda x_{t-1} + \cdots + \lambda^k x_{t-k} + \cdots \right) + \sum_{k=0}^{\infty} \lambda^k \varepsilon_{t-k}
\end{aligned}
$$

$$y_t = \alpha + \pi \left(x_t + \lambda x_{t-1} + \cdots + \lambda^k x_{t-k} + \cdots \right) + \sum_{k=0}^{\infty} \lambda^k \varepsilon_{t-k}$$

Then

$$\begin{aligned}
\frac{\partial y_t}{\partial x_t} &= \pi = \beta(1-\lambda) \\
\frac{\partial y_t}{\partial x_{t-1}} &= \frac{\partial y_{t+1}}{\partial x_t} = \pi\lambda < \pi \\
&\vdots \\
\frac{\partial y_t}{\partial x_{t-k}} &= \frac{\partial y_{t+k}}{\partial x_t} = \pi\lambda^k
\end{aligned}$$

Note

$$
\begin{aligned}
\lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial x_t} &= \lim_{k\to\infty} \pi\lambda^k = 0 \\
\sum_{k=0}^{\infty} \frac{\partial y_{t+k}}{\partial x_t} &= \pi(1+\lambda+\lambda^2+\cdots) = \frac{\pi}{(1-\lambda)} = \beta
\end{aligned}
$$

Another way to think about the long-run impact is to consider the unconditional mean for y_t :

$$\begin{aligned}
E[y_t] &= \gamma + \lambda E[y_{t-1}] + \pi E[x_t] = \gamma + \lambda E[y_t] + \pi E[x_t] \Rightarrow \\
(1-\lambda)\mu_y &= \gamma + \pi \mu_x \Rightarrow \\
\mu_y &= \frac{\gamma}{(1-\lambda)} + \frac{\pi}{(1-\lambda)}\mu_x = \frac{\alpha(1-\lambda)}{(1-\lambda)} + \frac{\beta(1-\lambda)}{(1-\lambda)}\mu_x \\
&= \alpha + \beta\mu_x
\end{aligned}$$

Then

$$\frac{\partial \mu_y}{\partial \mu_x} = \beta = \text{ long-run impact}$$

Autoregressive Distributed Lag (ADL) Model

$$\begin{aligned}
\phi(L)y_t &= c + \beta(L)x_t + \varepsilon_t \\
\phi(L) &= 1 - \phi_1 L - \cdots - \phi_p L^p \\
\beta(L) &= \beta_0 + \beta_1 L + \cdots + \beta_q L^q \\
x_t &= \text{strictly exogenous and is } I(0) \\
\text{roots of } \phi(z) &= 0 \text{ have modulus} > 1
\end{aligned}$$

Special case: ADL(1,1)

$$y_t = c + \phi y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

Q: How to interpret coefficients in ADL(1,1)?

Linear Regression with AR(1) Errors as Restricted ADL(1,1)

$$\begin{aligned} y_t &= \alpha + \beta x_t + u_t \\ u_t &= \phi u_{t-1} + \varepsilon_t, \ |\phi| < 1 \\ \text{Note} &: (1-\phi L)u_t = \varepsilon_t \end{aligned}$$

Multiply both sides by $(1-\phi L)$ and re-arrange

$$\begin{aligned} (1-\phi L)y_t &= (1-\phi L)\alpha + (1-\phi L)\beta x_t + (1-\phi L)u_t \\ &\Rightarrow y_t = (1-\phi)\alpha + \phi y_{t-1} + \beta x_t - \phi\beta x_{t-1} + \varepsilon_t \\ &= c + \phi y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \\ c &= \alpha(1-\phi) \\ \beta_0 &= \beta \\ \beta_1 &= -\phi\beta \end{aligned}$$