Formal definitions of endogenous, exogenous and predetermined variables

(A) In the context of a static or dynamic structural equation model, an <u>endogenous variable</u> y is one which is correlated with one or more disturbance terms in the model ie. $E[y_t u_s] \neq 0$ for some values of t and s.

(B) In the context of a static or dynamic structural equation model, an <u>exogenous variable</u> x is one which is <u>not</u> correlated with any disturbance terms in the model ie. $E[x_t u_s] = 0$ for all values of t and s.

(C) In the context of a dynamic structural equation model, a variable z is <u>predetermined</u> at time t if it is independent of all <u>current</u> and <u>future</u> disturbances in the model ie. $E[z_t u_s] = 0$ for all values of $s \geq t$.

To motivate these definitions, consider a simple dynamic Keynesian national income model:

$$C_t = \alpha + \beta Y_{t-1} + u_t$$
$$Y_t = C_t + I_t$$

To specify that I_t is an exogenous variable, an econometrician would simply write $E[I_t u_s] = 0$ for all t and s. To specify that Y_{t-1} is a predetermined variable at time t, an econometrician would write $E[Y_{t-1} u_s] = 0$ for $s \geq t$. The reduced form equations for the model would then be

$$C_t = \alpha + \beta Y_{t-1} + u_t$$
$$Y_t = \alpha + \beta Y_{t-1} + I_t + u_t$$

We can show formally that C_t and Y_t are both <u>current endogenous</u> variables by showing that $E[C_t u_t] \neq 0$ and $E[Y_t u_t] \neq 0$. Let $\sigma_u^2 \equiv V[u_t] = E[u_t^2]$ for all t. Then using the equations of the reduced form above we have

$$E[C_t u_t] = E[(\alpha + \beta Y_{t-1} + u_t)u_t]$$
$$= E[\alpha u_t + \beta Y_{t-1} u_t + u_t^2]$$
$$= \alpha E[u_t] + \beta E[Y_{t-1} u_t] + E[u_t^2]$$
$$= 0 + 0 + \sigma_u^2 = \sigma_u^2 \neq 0$$

Similarly,

$$E[Y_t u_t] = E[(\alpha + \beta Y_{t-1} + I_t + u_t)u_t]$$
$$= E[u_t^2] = \sigma_u^2 \neq 0$$

We can also show formally that Y_{t-1} is a lagged endogenous variable by showing that $E[Y_{t-1}u_{t-1}] \neq 0$. We deduce from the reduced form equation for income above that

$$Y_{t-1} = \alpha + \beta Y_{t-2} + I_{t-1} + u_{t-1}$$

Therefore

$$\begin{aligned}
E[Y_{t-1}u_{t-1}] &= E[(\alpha + \beta Y_{t-2} + I_{t-1} + u_{t-1})u_{t-1}] \\
&= E[\alpha u_{t-1} + \beta Y_{t-2}u_{t-1} + I_{t-1}u_{t-1} + u_{t-1}^2] \\
&= \alpha E[u_{t-1}] + \beta E[Y_{t-2}u_{t-1}] + E[I_{t-1}u_{t-1}] \\
&\qquad + E[u_{t-1}^2] \\
&= 0 + 0 + 0 + \sigma_u^2 = \sigma_u^2 \neq 0
\end{aligned}$$

Note that problems arise when the disturbance term is autocorrelated. If u_t is independent of its own past values $u_{t-1}, u_{t-2}, \ldots$, then although y_{t-1} depends on u_{t-1}, y_{t-1} can be independent of u_t and can then be regarded as predetermined. However, if u_t is correlated with u_{t-1}, then y_{t-1} and u_t must be correlated via their common link with u_{t-1}, and y_{t-1} cannot be predetermined.

This can easily be illustrated in the context of the Keynesian national income model above.

If $E[u_t u_{t-1}] = 0$, then Y_{t-1} is predetermined:

$$E[Y_{t-1} u_t] = E[(\alpha + \beta Y_{t-2} + I_{t-1} + u_{t-1}) u_t]$$
$$= E[\alpha u_t + \beta Y_{t-2} u_t + I_{t-1} u_t + u_{t-1} u_t]$$
$$= \alpha E[u_t] + \beta E[Y_{t-2} u_t] + E[I_{t-1} u_t] + E[u_{t-1} u_t]$$
$$= 0 + 0 + 0 + 0 = 0.$$

But suppose that u_t follows a first-order autoregressive process

$$u_t = \rho u_{t-1} + \varepsilon_t$$

where ε_t is "white noise". Then u_t is autocorrelated:

$$E[u_t u_{t-1}] = E[(\rho u_{t-1} + \varepsilon_t) u_{t-1}]$$
$$= E[\rho u_{t-1}^2 + \varepsilon_t u_{t-1}]$$
$$= \rho E[u_{t-1}^2] = \rho \sigma_u^2 \neq 0$$

In this case, Y_{t-1} <u>cannot</u> be predetermined at time t:

$$E[Y_{t-1} u_t] = E[(\alpha + \beta Y_{t-2} + I_{t-1} + u_{t-1}) u_t]$$
$$= E[\alpha u_t + \beta Y_{t-2} u_t + I_{t-1} u_t + u_{t-1} u_t]$$
$$= \alpha E[u_t] + \beta E[Y_{t-2} u_t] + E[I_{t-1} u_t] + E[u_{t-1} u_t]$$
$$= 0 + \beta E[Y_{t-2} u_t] + \rho \sigma_u^2 = \rho \sigma_u^2 \neq 0$$

So when u_t follows a first-order autoregressive process (and is therefore autocorrelated), Y_{t-1} <u>cannot</u> be predetermined. This creates problems for estimation.

(End of handout)