14.32 Recitation – Lagged Dependent Variables

Paul Schrimpf

May 1, 2009

- Wooldridge problems 10.1, 10.8, 11.8, 12.1, 12.2, 12.5

1 Lagged Dependent Variables

By a model with lagged dependent variables, we mean a regression with the lagged outcome included as an explanatory variable. The simplest example is:

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t \tag{1}$$

Sometimes you will hear the warning that if e_t is serially correlated, then $\hat{\beta}_1$ is inconsistent. Strictly speaking, this is true. For consistency, we require that $E[y_{t-1}e_t] = 0$ and if e_t and e_{t-1} are correlated, then this generally (but not necessarily) will not be true. However, Wooldridge and Prof Angrist both argue that it usually doesn't make sense to think of a model with lagged dependent variables as having serially correlated residuals (and I'd modify this statement to say that you shouldn't worry about serial correlation leading to inconsistency, but you should worry about it for your standard errors). This is another version of Josh's general point that OLS never gives inconsistent estimates; it may not give you consistent estimates of the regression you want, but it does give you consistent estimates of the regression you have. Put another way, OLS always gives you a consistent best linear approximation to the conditional expectation function. Whether or not the conditional expectation function is what you want depends on context. If you're interested in the causal effect of education or military status on earnings, then the conditional expectation function you can estimate probably isn't the conditinoal expectation function you want because you can't observe ability. However, it turns out that in many time series applications the regression you have is the regression you want.

To be more concrete about the math of lagged dependent variables, let's go over the example in Wooldridge section 12.1. So we estimate a regression that looks like (1),

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t$$

The question is whether $\hat{\beta}_1^{OLS}$ is consistent. The answer depends on what we have in mind. If we are interested in the best linear approximation to $E[y_t|y_{t-1}]$, say $y_t = \beta_0 + \beta_1 y_{t_1}$ where $\beta_1 = \frac{COV(y_{t-1}, y_t}{V(y_t)}$, then by construction $E[u_t y_{t-1}] = 0$ and OLS is consistent. Note that u_t might still be serially correlated, so we need to fix the standard errors, but at least OLS is consistent. If you are willing to go even further and say that $E[y_t|y_{t-1}, y_{t-2}, ...] = E[y_t|y_{t-1}]$[1] and that the CEF is linear, then you can also show that u_t is serially uncorrelated, so OLS standard errors are correct.

The situation when it makes sense to worry about serial correlation causing inconsistency is when you've derived (1) from some economic model so that β_0 and β_1 have an intrinsic meaning and are not necessarily coefficients of the best linear approximation to the conditional expectation function. Good examples of this situation are the (Neo-Keynesian) Philips curve (e.g. Galí and Gertler 1999), models about investment and

[1] Wooldridge calls this situation – when you've included all the variables relevant for the conditional expectation – **dynamic completeness**.

Tobin's Q (e.g. Blundell et al 1992), and consumption CAPM (e.g. Hansen and Singleton 1982). In these sort of models it's natural to think that u_t might be serially correlated, and as a result, correlated with y_{t-1}. For example if $u_t = \rho u_{t-1} + e_t$ where e_t is iid, then $E[u_t y_{t-1}] = \rho COV(y_{t-1}, u_{t-1}) = \frac{\rho\sigma_u^2}{1-\beta\rho} \neq 0$. In situations where β_0 and β_1 have intrinsic meaning as part of an economic model, the solution when you have serially correlated errors is to use instrumental variables. Ideally, the theory that led to your model will also suggest potential instruments. Often these instruments are simply older lags of the variables.

However, many time series applications are simply focused on estimating conditional expectation functions. In forecasting, the expectation of y_t given the past is the object of interest. Therefore, it doesn't make sense to worry about inconsistency due to serially correlated errors when forecasting. Another common situation in time series where the conditional expectation function is the object of interest is when estimating vector auto-regressions or VARs. In a VAR we have a handful of variables (say 2) that we're interested in the dynamics of. To summarize the dynamics we estimate:

$$\begin{aligned} y_t =& \beta_{yy1} y_{t-1} + ... + \beta_{yyp} y_{t-p} + \beta_{yx1} x_{t-1} + ... + \beta_{yxp} x_{t-p} + e_{yt} \\ x_t =& \beta_{xy1} y_{t-1} + ... + \beta_{xyp} y_{t-p} + \beta_{xx1} x_{t-1} + ... + \beta_{xxp} x_{t-p} + e_{xt} \end{aligned}$$

We then look at the estimates of β (or some function of them) either to: (1) gather some stylized facts that we think we should explain or (2) compare with the qualitative perdictions of some model. Here, we're not exactly forecasting, but all we care about is the dynamic behavior of y and x, that is, all we care about it is $E[y_t|y_{t-1}, x_{t-1}, ...]$ and $E[x_t|y_{t-1}, x_{t-1}, ...]$. [2]

[2] Sims (1980) first proposed VARs. A related idea are structural VARs. These try to impose weak restrictions that let us decompose movements in y and x into changes due to interpretable shocks. The classic example is Blanchard and Quah (1989) who decompose movements in GDP into those due to short-run (demand) shocks and long-run (productivity) shocks.