Dynamic Regression Models (Lect 15)

Ragnar Nymoen

University of Oslo

21 March 2013

- HGL: Ch 9; BN: Kap 10
- The HGL Ch 9 is a long chapter, and the testing for autocorrelation part we have already covered.
- HGL starts the chapter with the Finite Distributed lag model (DL), for example

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \varepsilon_t \qquad (1)$$

  and discuss estimation/testing with classical assumptions for ε_t, and without
- But (1) is "almost" a usual static model, and because economic relationships are often more genuinely dynamic, it has low practical relevance.
- Therefore we focus the "ARDL" part of Ch 9, and starts with the simplest version of that model class

Autoregressive first order model AR(1) model I

- The simplest "dynamic regression model". It has properties that carry over to more general models (ARDL below).
- Assume that we have $t = 1, 2, \ldots, T$ independent and identically distributed random variables ε_t:

$$\varepsilon_t \sim IID\left(0, \sigma_\varepsilon^2\right),\ t = 1, 2, \ldots, T$$

Then, from

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t, \quad |\beta_1| < 1, \quad \varepsilon_t \sim IID\left(0, \sigma_\varepsilon^2\right), \quad (2)$$

we know something precise about the *conditional* distribution of Y_t given Y_{t-1}, and more generally *the history of* Y *up to period* $t-1$.

Autoregressive first order model AR(1) model II

- $|\beta_1| < 1$ secures stationarity (HGL 9.1.3) for this model.
- We will refer to Y_t as given by (2) as a *1st order autoregressive process*, usually denoted AR(1).
- In direct parallel to the previous models we can write

$$Y_t = E(Y_t \mid Y_{t-1}) + \varepsilon_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t \qquad (3)$$

  where

$$E(\varepsilon_t Y_{t-1}) = 0 \qquad (4)$$

  by construction (in fact by assumption of $|\beta_1| < 1$, but leave that for another course)
- (4) is necessary for pre-determinedness of Y_{t-1}.
  - But is $E(\varepsilon_{t+j} Y_{t-1}) = 0$ for $j = 1, 2, \ldots$ as well?
  - And what about $E(\varepsilon_{t-1-j} Y_{t-1})$ for $j = 1, 2$?

Autoregressive first order model AR(1) model III

- To answer these questions: need to consider the solution of (2), which is a stochastic difference equation.

Solution I

- $|\beta_1| < 1$ defines Y_t as a *causal-process*: Stochastic shocks/impulses/news represented by ε come before (or in the same period) as the response in Y_t.
- The backward-recursive solution of a causal-process is dynamically stable. We show in class that it is:

$$Y_t = \beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t Y_0 + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{t-i} \tag{5}$$

where Y_0 is the *initial condition*.
The conditional expectation is

$$E(Y_t \mid Y_0) = \beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t Y_0$$

Solution II

while the **unconditional expectation of** Y_t is defined for the situation where $t \to \infty$:

$$E(Y_t) = \frac{\beta_0}{1-\beta_1} \tag{6}$$

For simplicity, we regard Y_0 as a deterministic parameter. Then the variance is found as:

$$\begin{aligned} Var(Y_t) &= Var(\sum_{i=0}^{t-1} \beta_1^i \varepsilon_{t-i}) = \sigma_\varepsilon^2 \sum_{i=0}^{t-1} \left(\beta_1^2\right)^i \\ &\underset{t\to\infty}{=} \frac{\sigma_\varepsilon^2}{1-\beta_1^2} \end{aligned} \tag{7}$$

Pre-determinedness of lagged Y I

The solution for Y_{t-1} (make use of (5)!) shows that:

$$E(Y_{t-1}\varepsilon_t) = E(\sum_{i=0}^{t-2} \beta_1^i \varepsilon_{t-i-1})\varepsilon_t = 0$$

and

$$E(Y_{t-1}\varepsilon_{t+j}) = 0 \text{ for } j = 1, 2, \ldots$$

But also that:

$$E(Y_{t-1}\varepsilon_{t-i}) \neq 0 \text{ for } i = 1, 2,$$

Y_{t-1} is a pre-determined explanatory variable.

Bias and consistency I

- To save notation: Consider the case of $E(Y_t) = 0 \Longrightarrow \beta_0 = 0$.
- The OLS estimator $\widehat{\beta}_1$ is

$$\widehat{\beta}_1 = \frac{\sum_{t=2}^{T} Y_t Y_{t-1}}{\sum_{t=2}^{T} Y_{t-1}^2} = \sum_{t=2}^{T}\left(\frac{\beta_1 Y_{t-1}^2}{\sum_{t=2}^{T} Y_{t-1}^2}\right) + \sum_{t=2}^{T}\left(\frac{Y_{t-1}\varepsilon_t}{\sum_{t=2}^{T} Y_{t-1}^2}\right) \quad (8)$$

$$\Longrightarrow$$

$$E\left(\widehat{\beta}_1 - \beta_1\right) = E\left(\frac{\sum_{t=2}^{T} Y_{t-1}\varepsilon_t}{\sum_{t=2}^{T} Y_{t-1}^2}\right)$$

- Cannot show that E of the bias term is zero

Bias and consistency II

- Both the denominator and numerator are random variables, and they are not independent: For example will ε_2 “be in” the numerator and (because of $Y_2 = \varepsilon_2$) also in $Y_2 \times Y_2$ in the denominator.
- But, with reference to the Law of large numbers and Slutsky's theorem we have

$$\text{plim}\left(\widehat{\phi}_1 - \phi_1\right) = \frac{\text{plim}\,\frac{1}{T}\sum_{t=2}^{T} Y_{t-1}\varepsilon_t}{\text{plim}\,\frac{1}{T}\sum_{t=2}^{T} Y_{t-1}^2} = \frac{0}{\frac{\sigma_\varepsilon^2}{1-\beta_1^2}} = 0.$$

since $E\left(Y_{t-1}\varepsilon_t\right) = 0$ (numerator) and $|\beta_1| < 1$ (implies the existence of the variance).

Bias and consistency III

- The OLS estimator $\widehat{\beta}_1$ in the AR(1) is consistent, and it can be shown to be asymptotically normal:

$$\sqrt{T}\left(\widehat{\beta}_1 - \beta_1\right) \xrightarrow{d} N\left(0, \left(1 - \beta_1^2\right)\right) \tag{9}$$

  which entails that *t-ratios* can be compared with critical values from the normal distribution.

- Therefore: the large sample inference theory for the regression model extends to the AR(1) model.

Analysis of finite sample bias in AR(1)

In (2), the finite sample bias can be shown to be approximately

$$E\left(\widehat{\beta}_1 - \beta_1\right) \approx \frac{-2\beta_1}{T},$$

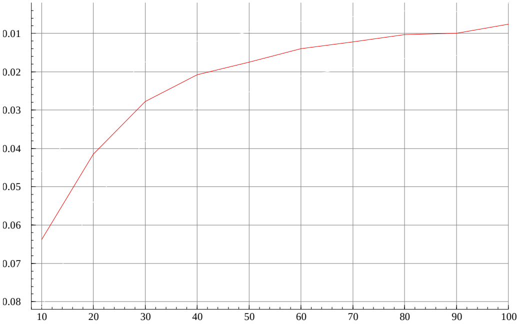
We can make this more concrete with a Monte-Carlo analysis.
In the experiment, the DGP is

$$Y_t = 0.5Y_{t-1} + \varepsilon_{Yt}, \quad \varepsilon_{Yt} \sim NIID\left(0, 1\right),$$

and $T = 10, 11, \ldots, 99, 100$. We use 1000 replications for each T and estimate the bias:

$$\hat{E}\left(\widehat{\beta}_{1(T)} - \beta_1\right) = \frac{1}{1000}\sum_{i=1}^{1000}\left(\widehat{\beta}_{1(T)i} - \beta_1\right).$$

Bias in the AR(1) model



$$\hat{E}\left(\widehat{\beta}_{1(10)} - 0.5\right) = -0.058 > \ldots \approx \frac{-2\times 0.5}{10} = -0.1$$

$$\hat{E}\left(\widehat{\beta}_{1(100)} - 0.5\right) = -0.008 > \ldots \approx \frac{-2\times 0.5}{100} = -0.01.$$

Monte Carlo analysis of AR(1) with exogenous regressor

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \varepsilon_t, \quad |\beta_1| < 1, \quad \varepsilon_t \sim IID\left(0, \sigma_\varepsilon^2\right). \tag{10}$$

which we will also refer to as an AutoRegressive Distributed Lag model, ARDL.

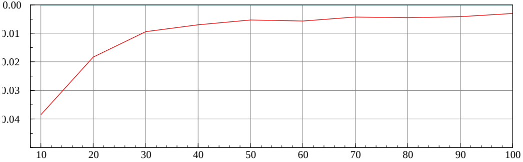
- We assume that X_t is stricty exogenous

Monte Carlo DGP:

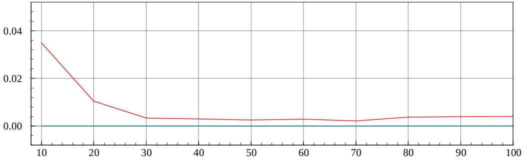
$$Y_t = 0.5 Y_{t-1} + 1 \cdot X_t + \varepsilon_{Yt}, \quad \varepsilon_{Yt} \sim NIID\left(0, 1\right),$$
$$X_t = 0.5 X_{t-1} + \varepsilon_{Xt}, \quad \varepsilon_{Xt} \sim NIID\left(0, 2\right),$$

There are now two biases, $\hat{E}\left(\hat{\beta}_{1(T)} - 0.5\right)$ and $\hat{E}\left(\hat{\beta}_{2(T)} - 1\right)$

Biases in the ADL model



$$\hat{E}\left(\hat{\beta}_{1(T)} - 0.5\right)$$



$$\hat{E}\left(\hat{\beta}_{2(T)} - 1\right)$$

Conclusions

- The OLS biases are small, and the speeds of convergence to zero are high
- OLS estimation, and the use $t-$ratios and F-statistics for testing extend to dynamic models, *given that the model is correctly specified*, disturbances that have the usual classical assumptions conditional on Y_{t-1} and X_t.
- In particular: Avoid residual autocorrelation because it will destroy pre-determinedness of Y_{t-1}!
- The tests we have covered for Non-Normality, Heteroskedasticity and Autocorrelation in (Lect 13 and 14) are valid mis-specification tests also for ARDL models!

Dynamic response to shocks

One purpose of estimating an ARDL model:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \varepsilon_t \tag{11}$$

with classical assumptions for ε_t conditional on

Y_{t-1}, X_t and X_{t-1}

is to estimate the dynamic response of Y to a permanent or temporary change in X.

- When we consider changes in the X, the key concept is **dynamic multiplier**.
- Can also study a temporary shock to ε (for example of magnitude one standard deviation σ) These dynamic effects are often called *impulse-responses.*
- In class: Derive dynamic multipliers (short), and show examples of estimated dynamic multipliers.
- Use of model in forecasting: Lecture 16.