Treating Time with All Due Seriousness

Professor Suzanna Linn[1]

[1]Department of Political Science
Penn State University

IMC: February 13, 2015

1. The GECM is appropriate when data are individually stationary or jointly stationary and cointegrated, that is when the equation is balanced.

1. The GECM is appropriate when data are individually stationary or jointly stationary and cointegrated, that is when the equation is balanced.
2. While fractional integration is conceptually appealing difficulties with estimation make it practically problematic as a modeling strategy in many cases.

1. The GECM is appropriate when data are individually stationary or jointly stationary and cointegrated, that is when the equation is balanced.
2. While fractional integration is conceptually appealing difficulties with estimation make it practically problematic as a modeling strategy in many cases.
3. Over fitting, a caution: don't ask too much of the data.

Goal of Applied Time Series Analysis

Goal of Applied Time Series Analysis

Estimate relationships among variables whose behaviors evolve over time and use estimates to test hypotheses and infer interesting features of the relationship in the long and short run.

Lesson #1: Choice of statistical model depends on temporal properties of the data

Stationary Time Series

When a stationary time series is shocked it returns relatively quickly to its long run mean.

$$Y_t = \pi Y_{t-1} + e_t$$
$$|\pi| < 1.0$$
$$e_t \sim (0, \sigma^2)$$

When the time series are all stationary, a variety of regression models can be used to model relationships.

Integrated (Unit Root) Time Series

When shocks to a series cumulate a series is integrated.

$$Y_t = Y_{t-1} + e_t$$
$$e_t \sim (0, \sigma^2)$$

Relationships between integrated time series can be modeled:

Integrated (Unit Root) Time Series

When shocks to a series cumulate a series is integrated.

$$Y_t = Y_{t-1} + e_t$$
$$e_t \sim (0, \sigma^2)$$

Relationships between integrated time series can be modeled:

- Using cointegration techniques if jointly stationary;
- Otherwise, the data must be made stationary through some form of transformation to avoid spurious inferences.

Fractionally Integrated Time Series

If shocks to a series decay hyperbolically a series is fractionally integrated.

Fractionally Integrated Time Series

If shocks to a series decay hyperbolically a series is fractionally integrated.

$$Y_t = \sum_{k=0}^{\infty} \pi_k Y_{t-k} + e_t$$
$$|\pi| < 1.0$$
$$e_t \sim (0, \sigma^2)$$

Relationships between fractionally integrated time series can be modeled:

Fractionally Integrated Time Series

If shocks to a series decay hyperbolically a series is fractionally integrated.

$$Y_t = \sum_{k=0}^{\infty} \pi_k Y_{t-k} + e_t$$

$$|\pi| < 1.0$$

$$e_t \sim (0, \sigma^2)$$

Relationships between fractionally integrated time series can be modeled:

- Using fractional cointegration techniques if jointly of a lower order of integration;
- Otherwise, the data must be made stationary through some form of transformation to avoid spurious inferences.

Lesson #2: Models must be balanced

The importance of balance

The importance of balance

- Stable long run relationships imply balanced equations.
- Intution: stable (stationary) stochastically bounded variables cannot cause – or be caused by – the path of a stochastically unbounded variable, the time series must diverge by ever larger amounts.
- Data must be transformed to ensure the left and (jointly the) righthand sides of the model are of equal orders of integration.

The GECM and Stationary Data

Consider the familiar ADL:

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{n} \sum_{i=0}^{q} \beta_j X_{jt-i} + \varepsilon_t$$

The assumptions:

- ε_t is white noise;
- $|\sum_{i=1}^{p} \alpha_i| < 1$ so that Y_t is stationary;
- the processes generating X_j are weakly exogenous for the parameters of interest such that $E(\varepsilon_t, X_{js}) = 0 \, \forall \, t, s$, and j;
- p refers to the number of lags of Y_t, q the number of lags of X_t, and n the number of exogenous regressors.

Let $p = q = n = 1$:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$$

- Short run effects (impact multipliers) are given by β_0 and β_1.
- Long run effects are given by $\frac{\beta_0+\beta_1}{1-\alpha_1}$.
- The long run equilibrium is given by $Y^* = \frac{\alpha_0}{1-\alpha_1} + \frac{\beta_0+\beta_1}{1-\alpha_1} X^*$.
- The error correction rate is given by $(\alpha_1 - 1)$.

Equivalence of the GECM

Class of models whose general form is equivalent to the ADL.

- First subtract Y_{t-1} from both sides of the ADL.

$$\Delta Y_t = \alpha_0 + (\alpha_1 - 1)Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$$

- Add and subtract $\beta_0 X_{t-1}$ on the right side.

$$\Delta Y_t = \alpha_0 + (\alpha_1 - 1)Y_{t-1} + \beta_0 \Delta X_t + (\beta_0 + \beta_1)X_{t-1} + \varepsilon_t$$

- If we stop here we can write the generalized ECM (GECM) as

$$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t.$$

By substitution, we see the equivalence: $\alpha_1^* = (\alpha_1 - 1)$, $\beta_0^* = \beta_0$, and $\beta_1^* = \beta_0 + \beta_1$ (De Boef and Keele 2008).

Grant and Lebo maintain the GECM cannot be used for stationary time series.

If the error correction rate, $\alpha_1^* = -1.0$ and $\alpha_1^* = \alpha_1 - 1$, then $\alpha_1 = 0$:

In the ADL:

$$Y_t = \alpha_0 + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$$

and LRM$= \frac{\beta_0+\beta_1}{1-\alpha_1} = \beta_0 + \beta_1$.

In the GECM:

$$\Delta Y_t = \alpha_0 - 1.0 Y_{t-1} + \beta_0^* X_t + \beta_1^* X_{t-1} + \varepsilon_{GECM,t}.$$

and $\text{LRM}_{GECM} = -\frac{\beta_1^*}{\alpha_1^*} = \beta_1^* = \beta_0 + \beta_1$.

- $\rightarrow$ The LRM in the distributed lag model equals that in the GECM.
- $\rightarrow$ Y returns to its long run value after one period.
- $\rightarrow$ If $\beta_0 = \beta_1 = 0$ then $\beta_1^* = 0$ and there exists no long run relationship between X and Y even with significant error correction.

Some simulations

We simulated:

$$\begin{aligned} Y_t &= 0.5Y_{t-1} + e_{1t} \\ X_t &= \rho X_{t-1} + e_{2t} \\ e_{1t}, e_{2t} &\sim IID. \end{aligned}$$

We estimated the GECM.

$$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t.$$

The true parameters of the GECM are:

- $\alpha_1^* = -0.50$,
- $\beta_0^* = 0.0$,
- $\beta_1^* = 0.0$, and
- LRM $= -\frac{\beta_1^*}{\alpha 1^*} = 0$.

The GECM with 2 Unrelated Autoregressive Series

Rejection Rates on Estimated Coefficients

T	$\hat{\rho}$	$\hat{\alpha_1^*}$	$\hat{\beta}_1^*$	$\hat{\beta}_1^* / -\hat{\alpha}_1^*$
100	0	1.000	0.076	0.067
100	0.1	1.000	0.051	0.044
100	0.5	1.000	0.056	0.051
100	0.7	1.000	0.052	0.052
100	0.9	1.000	0.082	0.094

The GECM with 2 Unrelated Autoregressive Series

Bias in Estimated Coefficients

T	$\hat{\rho}$	$\hat{\alpha_0}$	$\hat{\alpha_1^*}$	$\hat{\beta_0^*}$	$\hat{\beta_1^*}$	$\hat{\beta}_1^* / -\hat{\alpha}_1^*$
100	0	-0.003	-0.024	0.000	-0.003	0.007
100	0.1	0.002	-0.027	0.003	0.004	-0.009
100	0.5	0.005	-0.034	0.000	-0.004	0.009
100	0.7	-0.004	-0.034	-0.003	0.002	-0.003
100	0.9	0.002	-0.041	-0.003	-0.001	0.002

There is no evidence the analyst would conclude that there is a long run relationship between X and Y at unacceptable levels.

What happens if X and Y are causally related?

We simulated:

$$Y_t = 0.5Y_{t-1} + 0.25X_t + 0.50X_{t-1} + e_{1t}$$
$$X_t = \rho X_{t-1} + e_{2t}.$$

We estimated the GECM.

$$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t.$$

Following the algebraic transformation, the true parameters of the GECM are:

- $\alpha_1^* = \alpha_1 - 1 = -0.50$,
- $\beta_0^* = \beta_0 = 0.25$,
- $\beta_1^* = \beta_0 + \beta_1 = 0.75$, and
- LRM $= -\frac{\beta_1^*}{\alpha 1^*} = \frac{\beta_0+\beta_1}{1-\alpha_1} = 1.50$.

The GECM with 2 Related Autoregressive Series

Bias In Estimated Coefficients

T	$\hat{\rho}$	$\hat{\alpha_0}$	$\hat{\alpha_1^*}$	$\hat{\beta_0^*}$	$\hat{\beta_1^*}$	$\hat{\beta}_1^*/-\hat{\alpha}_1^*$
100	0	0.002	-0.016	-0.003	-0.004	0.026
100	0.1	-0.002	-0.022	0.002	0.008	0.022
100	0.5	0.003	-0.017	0.000	0.015	0.006
100	0.7	-0.002	-0.018	-0.006	0.010	0.020
100	0.9	-0.002	-0.018	-0.001	0.018	0.011

- Bias on effects of X_{t-1} average (across 80 experiments) 0.012 or about 1.5% of the true value.
- Bias on the LRM averages 0.016 or about 1% the true value.

The GECM with 2 Related Autoregressive Series

Rejection Rates on Estimated Coefficients

T	$\hat{\rho}$	$\hat{\alpha}_1^*$	$\hat{\beta}_0^*$	$\hat{\beta}_1^*$	$\hat{\beta}_1^*/-\hat{\alpha}_1^*$
100	0	1.000	1.000	1.000	0.935
100	0.1	1.000	0.999	0.999	0.938
100	0.5	1.000	1.000	1.000	0.947
100	0.7	1.000	1.000	1.000	0.944
100	0.9	1.000	1.000	1.000	0.951

The GECM is appropriate for stationary data

A word of caution

Logic of Error Correction

- Anytime X and Y are out of equilibrium $Y \neq \frac{\alpha_0}{1-\alpha_1} + \frac{\beta_0+\beta_1}{1-\alpha_1}X$, Y must adjust to return the series to equilibrium.
- Call the equilibrium error μ_t. If at time $t-1$ μ_t is positive, the value of Y_t is too high (above its equilibrium value), Y must adjust downward at time t. If μ_t is negative, Y_t must adjust upward.
- Movement in Y_t is in the opposite direction of the disequilibrium $\rightarrow$ the error correction rate must be negative.

Error Correction Rates and Model (Mis)Specification

Error Correction Rates and Long Run Equilibria

α_1^*	α_1	Diagnosis
$0 > \alpha_1^* > -1.0$	$0 < \alpha_1 < 1.0$	Steady return to long run equilibrium.
$-2.0 < \alpha_1^* < -1.0$	$-1.0 < \alpha_1 < 0$	Oscillating return to long run equilibrium.
$\alpha_1^* > 0$	$\alpha_1 > 1.0$	no long run equilibrium exists.
$\alpha_1^* < -2.0$	$\alpha_1 < -1.0$	no long run equilibrium exists.
$\alpha_1^* = 0$	$\alpha_1 = 1.0$	no long run equilibrium exists.

Fractional Integration?

Fractional Integration

The general ARFIMA(p, d, q) model is given by:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right) (1 - L)^d Y_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t$$

where

- p refers to the number of autoregressive parameters, ϕ,
- q refers to the number of moving average parameters, θ,
- and d is the fractional differencing parameter.

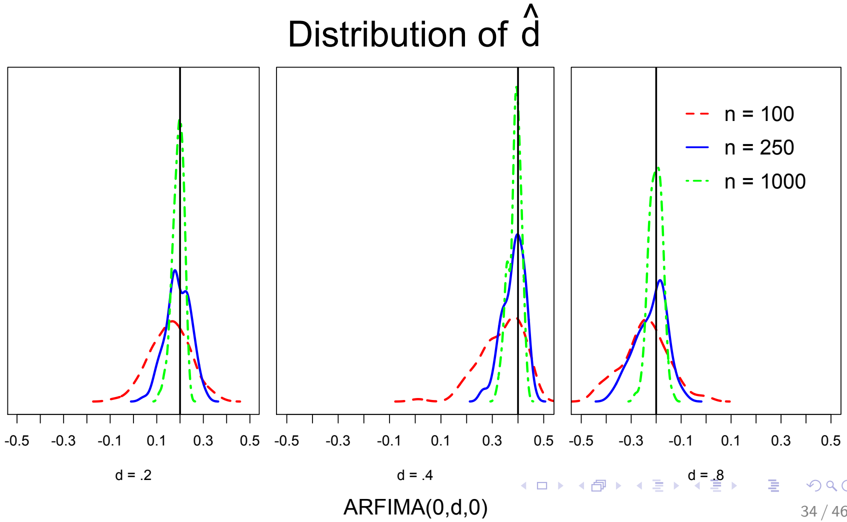
Potential Problems with ARFIMA Models

- Tests often suggest the data is fractionally integrated when it's not (Diebold & Inoue 2001; Engle & Smith 1991Granger & Hyung 1999).
- Bhardwaj & Swanson (2003, 2006) show:
  - spurious long memory often arises in tests of short memory;
  - standard tests produce evidence for long-memory even where predictions from a number of ARFIMA models fare worse from those with more standard models;
  - under some circumstances ARFIMA models offer superior predictions to alternative models about half the time but only when the sample sizes are large and forecast horizons long.
- Granger (1999) noted that ARFIMA models may fall into an "empty box".

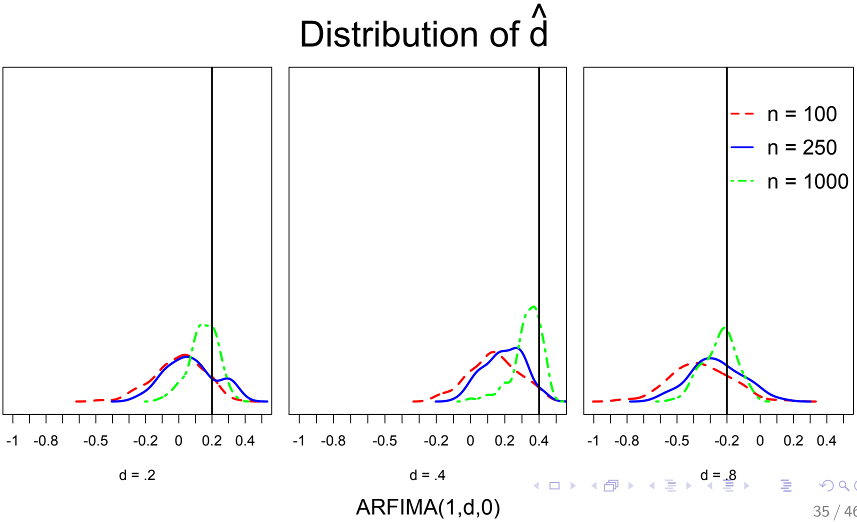
Can we reliably estimate d?

- We simulated the ARFIMA process given above for samples of size 50, 100, 250, 500, 1000, and 1500.
- We allowed for a range of dynamics, including ARFIMA(0,d,0), ARFIMA$(1, d, 0)$, ARFIMA$(0, d, 1)$ and ARFIMA$(1, d, 1)$.
- d takes on values 0 (no fractional integration), 0.20, 0.40, 0.45, and 0.80. (In the latter case the data is integer differenced before simulation and estimation so that $d = -0.20$ in the transformed data).
- We estimate the ARFIMA process under the optimal, but unrealistic assumption the order of the short run dynamics is known. Assuming otherwise increases are uncertainty that the selected process mimics the DGP.

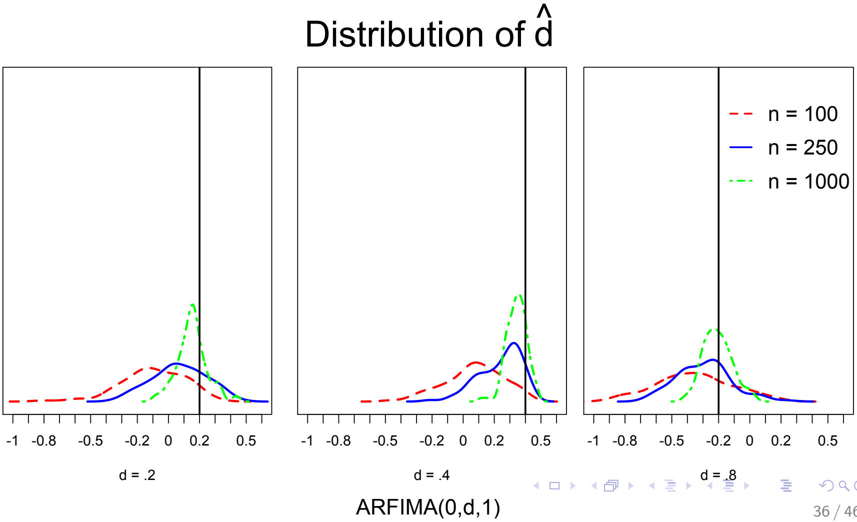
Can we reliably estimate d?



Can we reliably estimate d?



Can we reliably estimate d?



Can we reliably estimate d?

Summary

- Exact MLE produces downwardly biased estimates of d in all models, especially when $t = 100$ and $t = 250$ and is worse when the DGP contains short run dynamics.
- The estimator has particular problem distinguishing long from short run dynamics.
- In some cases estimates range across almost all possible values of d.

What happens when there is no fractional integration?

We over diagnose fractional integration when $d = 0$:

Model	t	Mean	Rejection Rate
ARMA(0,0)	100	$-.032$	11 %
	250	$-.017$	9 %
	1,000	$-.004$	13 %
ARMA(1,0)	100	$-.199$	32 %
	250	$-.121$	34 %
	1,000	$-.073$	21 %
ARMA(0,1)	100	$-.227$	16 %
	250	$-.132$	34 %
	1,000	$-.019$	12 %
ARMA(1,1)	100	$-.530$	69 %
	250	$-.316$	53 %
	1,000	$-.056$	27 %

While fractional integration is conceptually appealing difficulties with estimation make it practically problematic as a modeling strategy in many cases.

Overfitting

- Rule of thumb: one should fit one parameter for each 10 observations when the data are independent and identically distributed (Babyak 2004).
- When there is not enough information in the data, the model can be tuned to fit random patterns instead of to the conditional expectation which is generally of interest.
- The likelihood of finding spurious relationships is quite high when models are over fit.

Article	Time Periods	# of Parameters
Casillas, Enns, Wohlfarth (2011)	45	7
Ura and Ellis (2012)	36	11
Sanchez et al. (2011)	60	11
Kelly and Enns (2010)	54	8
Volscho and Kelly (2012)	60	10

The Power of Statistical Tests

Keele and Kelly (2006)

Asked: How long does a series need to be to reliably detect autocorrelation in models with lagged dependent variables?

- They found the analyst needs sample sizes between 250 and 500 observations before these tests had much power.
- The implication is that these tests may not help us diagnose model performance in many applications.

Where does this leave us?

Treating Time with All Due Seriousness

1. Do not ask too much of small samples.
   - Be wary of over fitting
   - Know that estimate of d are highly uncertain.
2. Make sure your equations are balanced.
3. Heed the clues implied by estimated error correction rates near outside the permissible bounds.