Modeling Interactions in Error Correction Models

Flávio D. S. Souza
fsouza@tamu.edu
people.tamu.edu/~fsouza

Abstract

Error correction models (EMCs) are the preferred modeling choice when cointegration is present (Engle & Granger 1987). But ECMs are also relatively inflexible and econometrically complex. This complicates modeling the effects of interactions within ECMs. In this paper, I evaluate the implications of including a multiplicative interaction within ECMs and suggest a strategy to do so correctly. I discuss the trade-offs between modeling strategies using a series of Monte Carlo simulations. I find that omitting interactions within ECMs, when the true DGP includes one, leads to biased and inefficient parameter estimates. I conclude with suggestions for future research on this topic.

Background

- A series with a unit root (we also call this a random walk) is integrated of order one, I(1)
- A stationary series with no trend (as well as an I(1) series that has been differenced once) is said to be integrated of order 0, I(0)
- Cointegration occurs when two series share common stochastic trends (Stock and Watson 1988; Enders 2010)
- It is suggested that we use an error correction technique to model cointegrated data
- If all *x*'s were I(1), this would be:

$$\Delta^p y_t = \alpha_0 + \sum_{f=1}^{p} \alpha_f t_{t-p} + \sum_{g=1}^{q}\sum_{h=0}^{r} \beta_{g,h}\Delta x_{g,t-h} + \sum_{g=1}^{q}\sum_{h=0}^{r} \gamma_{g,h} x_{g,t-h-1} + \varepsilon_t$$

where *p*, *q*, and *r* come from the equivalent ARDL(*p*, *q*, *r*) model and stand for the number of lags of *y*, number of lags of *x*, and the number of exogenous regressors, respectively. I assume $\varepsilon_t \sim N(0,\sigma^2)$.

- In the case where *p*, *q*, and *r* are each equal to 1, the last equation simplifies to

$$\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 \Delta x_t + \gamma_1 x_{t-1} + \varepsilon_t$$

- Arnade, Kuchler, and Calvin (2011), emphasize that ECMs are composed of equilibrium and disequilibrium components.
- The equilibrium component includes all of the lagged terms and the disequilibrium component includes all of the differenced terms
  - This is the equilibrium equation: $y_t = \alpha + \beta_1 x_t + \mu_t$
  - This is the disequilibrium equation: $\Delta y_t = \beta_2 \Delta x_t + \varepsilon_t$

How Have Previous Scholars Dealt With Interactions in ECMs?

- Ignored the interaction
- Resorted to sample splitting (see Philips, Souza, and Whitten 2020)
- Modeled them in various ways (as reflected in Zack Warner's *Working Paper*):

Warner's Least General

(I) $\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 \Delta x_t + \beta_1 x_{t-1} + \beta_2 \Delta z_t + \beta_3 z_{t-1} + \varepsilon_t$

This is is a traditional ECM with two covariates, assuming all variables are I(1)

⋮

(II) $\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 \Delta x_t + \beta_1 x_{t-1} + \beta_2 \Delta z_t + \beta_3 z_{t-1} + \beta_4 \Delta x_t z_{t-1} + \varepsilon_t$

⋮

Here are some examples of ECMs assuming an interactive relationship but restricting some other possible interaction terms

(III) $\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 \Delta x_t + \beta_1 x_{t-1} + \beta_2 \Delta z_t + \beta_3 z_{t-1} + \beta_6 x_{t-1} z_{t-1} + \beta_7 \Delta x_t \Delta z_t + \varepsilon_t$

⋮

(IV) $\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 \Delta x_t + \beta_1 x_{t-1} + \beta_2 \Delta z_t + \beta_3 z_{t-1} + \beta_4 \Delta x_t z_{t-1} + \beta_5 x_{t-1} \Delta z_t + \beta_6 x_{t-1} z_{t-1} + \beta_7 \Delta x_t \Delta z_t + \varepsilon_t$

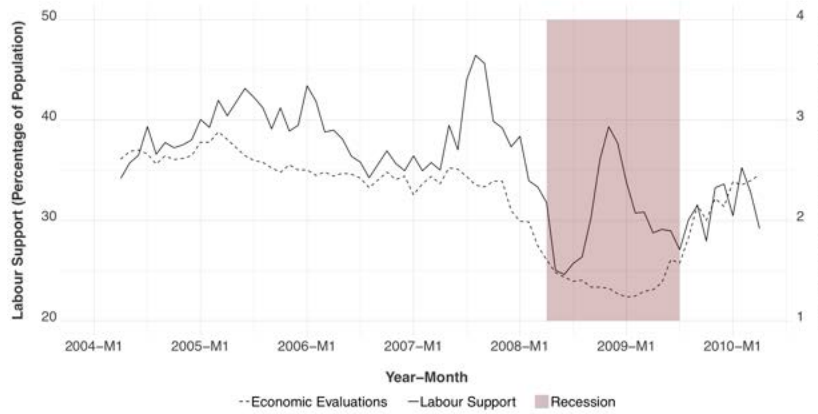
This is Warner's "most general" ECM, assuming all variables are I(1)

Warner's Most General

Issues

- But what should we do if we believe that all of the x's in the model interact with another exogenous variable, say z_t?
- What about contexts where we believe that the adjustment parameter is conditional on values of an exogenous right-hand-side variable?

For example:



How Should We Deal with These Issues?

If we believe that the rate of cointegration is itself conditional on values of an exogenous variable, then we would interact that exogenous variable by the entire equilibrium component of the model

- In my paper, I focus on interactions between the equilibrium component and a dummy variable:

$$\Delta y_t = \gamma_1 (y_{t-1} - \alpha - \beta_1 x_{t-1}) \times (1 + \gamma_2 z_t) + \beta_2 \Delta x_t + \varepsilon_t$$

What Happens If We Ignore This Interaction?

If the true DGP contains this interactions,

- We will recover an incorrect adjustment parameter
- We will obtain a single adjustment parameter, when in reality it varies depending on values of z_t
- We will estimate long-term parameters incorrectly
- The model will suffer from omitted variable bias

Monte Carlo Simulations

I ran 1000 Monte Carlo simulations (with 50 observations each) to explore what happens when we estimate the wrong model under the following four DGPs:

(1) No interaction:

$$\Delta y_t = -0.25 y_{t-1} + 4\Delta x_t + 8 x_{t-1} + \varepsilon_t$$

(2) Controlling for z_t (and, still, no interaction):

$$\Delta y_t = -0.25 y_{t-1} + 4\Delta x_t + 8 x_{t-1} + 2 z_t + \varepsilon_t$$

(3) Interaction between: z_t and $\Delta x_t, x_{t-1}$.

$$\Delta y_t = -0.25 y_{t-1} + 4\Delta x_t + 8 x_{t-1} + 3 x_{t-1} z_t + 1 \Delta x_t z_t + 2 z_t + \varepsilon_t$$

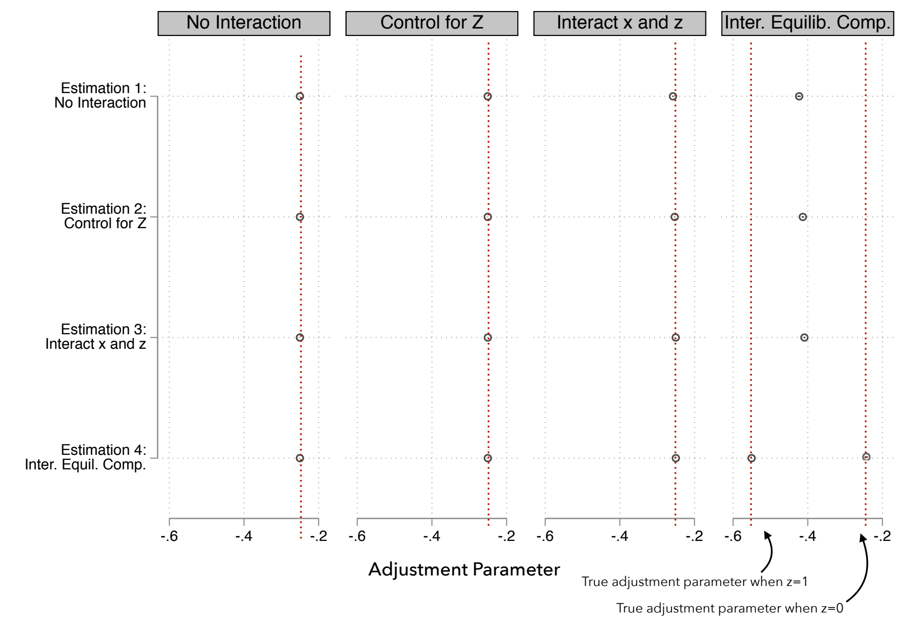
(4) Interaction between: z_t and the entire equilibrium component of the model

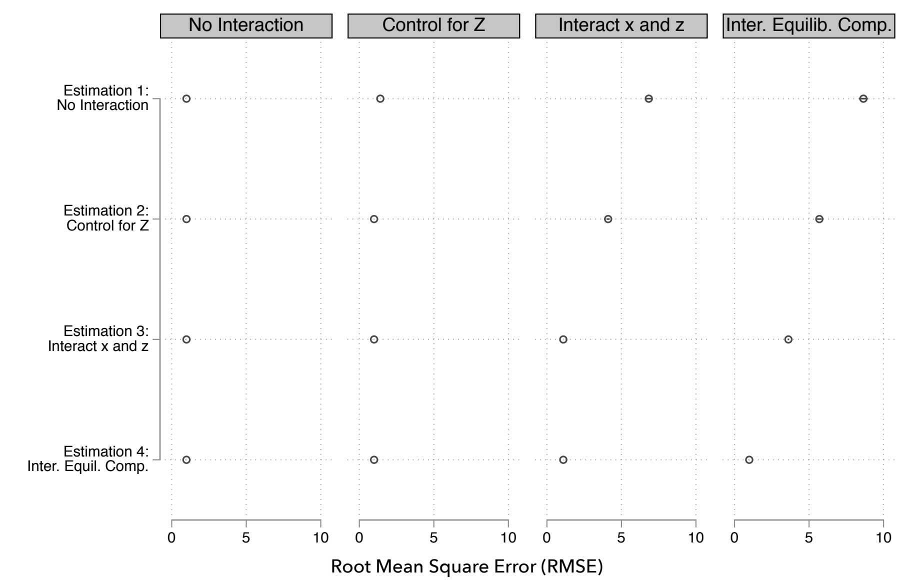
$$\Delta y_t = -0.25 y_{t-1} + 4\Delta x_t + 8 x_{t-1} + 3 x_{t-1} z_t + 0.3 y_{t-1} z_t + 2 z_t + \varepsilon_t$$

where $\varepsilon_t \sim N(0,\sigma^2)$.

Simulation Results

Note: each column represents a separate DGP; each row represents an estimation strategy. Red vertical lines represent true parameter values. Notice two adjustment parameters for the fourth DGP below. One is for a scenario where z=1, the other for a scenario where z=0.





Conclusion

- Results suggest that little bias is incurred when estimating an interaction between an exogenous variable and the equilibrium component if the interaction is not in the DGP;
- But omitting an interaction between an exogenous variable and the equilibrium component, when the true DGP includes one, leads to a strong bias in the adjustment parameter estimate;
- There are major efficiency gains to estimating the correct model in every case; but the efficiency cost of estimating an interaction between an exogenous variable and the equilibrium component of the model is marginal when no interaction (or a different interaction) is present.
- Future work: additional simulations are needed with other specifications, as well as under scenarios where the explanatory variables are correlated with one another.