Nonlinear Cointegrating Regression under Weak Identification

Peter Phillips and Xiaoxia Shi

Yale University and University of Wisconsin - Madison

Oct. 2010

Nonlinear Cointegrating Regression

- We deal with the nonlinear regression model:

$$Y_t = \beta_0 g(X_t, \pi_0) + \varepsilon_t.$$

Nonlinear Cointegrating Regression

- We deal with the nonlinear regression model:

$$Y_t = \beta_0 g(X_t, \pi_0) + \varepsilon_t.$$

  - g is known up to parameter π_0,

Nonlinear Cointegrating Regression

- We deal with the nonlinear regression model:

$$Y_t = \beta_0 g(X_t, \pi_0) + \varepsilon_t.$$

  - g is known up to parameter π_0,
  - X_t is $I(1)$: $X_t = X_{t-1} + v_t$, v_t stationary.

Nonlinear Cointegrating Regression

- We deal with the nonlinear regression model:

$$Y_t = \beta_0 g(X_t, \pi_0) + \varepsilon_t.$$

  - g is known up to parameter π_0,
  - X_t is $I(1)$: $X_t = X_{t-1} + v_t$, v_t stationary.
  - ε_t: martingale difference sequence.

Nonlinear Cointegrating Regression

- We deal with the nonlinear regression model:

$$Y_t = \beta_0 g(X_t, \pi_0) + \varepsilon_t.$$

  - g is known up to parameter π_0,
  - X_t is $I(1)$: $X_t = X_{t-1} + v_t$, v_t stationary.
  - ε_t: martingale difference sequence.
  - β_0: loading coefficient, unknown.

Nonlinear Cointegrating Regression

- We deal with the nonlinear regression model:

$$Y_t = \beta_0 g(X_t, \pi_0) + \varepsilon_t.$$

  - g is known up to parameter π_0,
  - X_t is $I(1)$: $X_t = X_{t-1} + v_t$, v_t stationary.
  - ε_t: martingale difference sequence.
  - β_0: loading coefficient, unknown.
- Y_t may or may not be $I(1)$: nonlinear transform relates variables with different memory status.

Weak Identification

- If $\beta_0 = 0$, π_0 is not identified
- If β_0 is close to zero, π_0 is weakly identified.
- β_0 close to zero: Y_t is local to a martingale difference sequence.
- A potentially good model for stock market returns, as Y_t is only weakly predictable by an integrated X_t.

What's the Problem?

- Park and Phillips (2001) study NLS estimators of such a model assuming strong ID.
- Their asymptotic results are misleading under weak ID.
- Extreme case: $\beta_0 = 0$:
  - 'π_0 is not identified,
  - Nonlinear least square estimator $\hat{\pi}_n$ should be no where near π_0.
  - But strong ID asymptotics would suggest: $\hat{\pi}_n \rightarrow_p \pi_0$.

What's the Problem?

- In a less extreme case, β_0 is close to zero.
- π_0 is weakly identified.
- The exact distribution of $\hat{\pi}_n$ is not approximated well by the asymptotic distributions derived under strong ID.
- Not only $\hat{\pi}_n$ is affected. $\hat{\beta}_n$ is also affected because the estimation of β depends on the quality of the estimation of π_0.
- Confidence intervals based on the strong ID asymptotic theory are useless under weak ID.

Contributions

- Develop a new local limit theory that approximate the distributions of $\hat{\pi}_n$ and $\hat{\beta}_n$ uniformly well regardless of identification strength.
- Construct weak identification-robust confidence intervals.
- Show that the minimum coverage probability of those robust confidence intervals converges to the nominal level.
- Uniform weak convergence of (properly scaled) $\sum_{t=1}^{n} g(X_t, \pi) u_t$ is proved on the way.

Local Limit Theory

- Let the true parameters to drift with sample size: denote the true parameter sequence by β_n
- Allow β_n to drift to zero.
- By doing this, we can mimic the weak identification situation.
- Derive the asymptotic distribution of the properly scaled and centered NLS estimators under the drifting sequences.

Nonlinear Least Square

- We consider the NLS estimators:

$$\begin{aligned} (\hat{\beta}_n, \hat{\pi}_n) &= \arg\min_{\beta,\pi} Q_n(\beta, \pi) \\ Q_n(\beta, \pi) &= n^{-1} \textstyle\sum_{t=1}^n (Y_t - \beta g(X_t, \pi))^2. \end{aligned}$$

- Concentrate π out:

$$Q_n(\pi) = n^{-1} \textstyle\sum_{t=1}^n \left(Y_t - \hat{\beta}_n(\pi) g(X_t, \pi)\right)^2,$$

where

$$\hat{\beta}_n(\pi) = \frac{\sum_{t=1}^n Y_t g(X_t, \pi)}{\sum_{t=1}^n g^2(X_t, \pi)}$$

Consistency and Asymptotic Distributions

- It turns out:

$$\left[Q_n(\pi) - n^{-1}\sum_{t=1}^{n} Y_t^2\right] = -\frac{n^{-1}\left(\sum_{t=1}^{n} Y_t g(X_t,\pi)\right)^2}{\sum_{t=1}^{n} g^2(X_t,\pi)}$$

$$= \quad \frac{n^{-1}\left(\sum_{t=1}^{n} u_t g(X_t,\pi) + \beta_n \sum_{t=1}^{n} g(X_t,\pi_n) g(X_t,\pi)\right)^2}{\sum_{t=1}^{n} g^2(X_t,\pi)}$$

- If β_n drifts to zero slow, the second term dominates, $\hat{\pi}_n \rightarrow_p \pi_0$.
- If β_n drifts to zero fast, the second term does not dominate, $\hat{\pi}_n$ has random limit.

Consistency and Asymptotic Distributions

- Asymptotic distributions of $\hat{\pi}_n - \pi_0$ (properly scaled) are derived with more tedious algebra.
- So are the consistency and asymptotic distribution of $\hat{\beta}_n$.
- Note that I haven't told you what counts as "slow" or "fast".

Drifting Rates

$$\frac{n^{-1}\left(\sum_{t=1}^{n} u_t g(X_t,\pi) + \beta_n \sum_{t=1}^{n} g(X_t,\pi_n) g(X_t,\pi)\right)^2}{\sum_{t=1}^{n} g^2(X_t,\pi)}$$

- The threshold drifting rates of β_n depends on the convergence rates of $\sum_{t=1}^{n} u_t g(X_t,\pi)$ and $\sum_{t=1}^{n} g(X_t,\pi_n) g(X_t,\pi)$.
- Those depend on the functional form of g,
- because X_t is $I(1)$.
- This differs from the stationary case (Andrews and Cheng (2010))

Drifting Rates

- We study two types of functions g.
- Integrable functions: $\int g(x,\pi)dx < \infty$.
  - $n^{-1/4}\sum_{t=1}^{n} u_t g(X_t,\pi) \to_d L(1,0)^{1/2} Z(\pi)$
  - $n^{-1/2}\sum_{t=1}^{n} g(X_t,\pi_n)g(X_t,\pi) \to_d L(0,1)\int g(x,\pi_n)g(x,\pi)dx$
  - Therefore, $n^{-1/4}$ is the threshold drifting rate.
- $n^{1/4}\beta_n \to \infty$: "slow" "near strong ID"
- $n^{1/4}\beta_n \to c < \infty$: "fast" "weak ID"

Drifting Rates

- Asymptotically Homogeneous functions: functions that are asymptotically equivalent to x^{κ}.
  - Asymptotic order usually depends on π
  - Convergence rates of those covariance processes depend on the asymptotic order
- Thus, the threshold drifting rate of β_n depends on the asymptotic order of g (and through which depends on π)

Confidence Intervals

- The asymptotic distributions of $\hat{\pi}_n$ are rather tedious looking, but we can characterize it in the following way:

$$d_n \hat{\beta}_n (\hat{\pi}_n - \pi_0) \rightarrow_d \begin{cases} T(c, \pi_0) \text{ if } d_n \beta_n \rightarrow c \in R \\ T(\infty, \pi_0) \text{ if } d_n \left| \beta_n \right| \rightarrow \infty. \end{cases}$$

Thus, the critical value for a robust C.I. can be:

$$\hat{q} = \begin{cases} \sup_c q_{1-\alpha} \left(\left| T(c, \hat{\pi}_n) \right| \right) \text{ if } k_n \left| \hat{\beta}_n \right| \leq 1 \\ q_{1-\alpha} \left(\left| T(\infty, \hat{\pi}_n) \right| \right) \text{ if } k_n \left| \hat{\beta}_n \right| > 1. \end{cases}$$

where $k_n = o(d_n)$.

Confidence Intervals

- The confidence interval is asymptotically valid:

$$\inf_{\beta,\pi} \Pr_{\beta,\pi} \left(\left| d_n \hat{\beta}_n (\hat{\pi}_n - \pi_0) \right| \leq \hat{q} \right) \to \alpha.$$

- Robust confidence intervals for β maybe constructed in a similar fashion.

A Technical Contribution

- We show for integrable functions:

$$n^{-1/4} \sum_{t=1}^{n} u_t g(X_t, \cdot) \Longrightarrow L(1,0)^{1/2} Z(\cdot),$$

where $Z(\cdot)$ is a Gaussian process with covariance kernel $\sigma_u^2 \int g(x,\pi) g(x,\pi') dx$.

- We show for asymptotically homogeneous functions:

$$n^{-1/2} \kappa_n^{-1}(\cdot) \sum_{t=1}^{n} u_t g(X_t, \cdot) \Longrightarrow \int h(V, \cdot)\, dU,$$

where $\kappa_n(\cdot)$ is the asymptotic order of the function, (U, V) is a vector Brownian motion.

Summary

- We developed a local limit theory for NLS estimators of the weakly identified cointegrating regression model.
- We constructed weak-identification-robust confidence intervals for the parameters based on the local limit theory.
- We showed the uniform weak convergence of the covariance process between $g(X_t, \pi)$ and u_t, strengthening the results in Park and Phillips (2001).