Discussion of
"Many Terms in a Series Estimator of the Partially Linear Model"

by **M. Cattaneo**, M. Jansson, and W. Newey

Alfonso Flores-Lagunes
University of Florida

4th Annual Meeting of IEN-LACEA
Miami FL, October 2010

Contributions and Message of the Paper

1. New asymptotic approximation to the distribution of the series-based partially linear model (PLM)

Contributions and Message of the Paper

1. New asymptotic approximation to the distribution of the series-based partially linear model (PLM)
   - Accounts for the potentially large number of series-based terms→ more precise approximation

Contributions and Message of the Paper

1. New asymptotic approximation to the distribution of the series-based partially linear model (PLM)
    - Accounts for the potentially large number of series-based terms→ more precise approximation
2. Identifies a problem with the variance of this estimator under “classical” asymptotics

Contributions and Message of the Paper

1. New asymptotic approximation to the distribution of the series-based partially linear model (PLM)
   - Accounts for the potentially large number of series-based terms→ more precise approximation
2. Identifies a problem with the variance of this estimator under "classical" asymptotics
   - Importantly, this problem also permeates to OLS with many covariates!!

Contributions and Message of the Paper

1. New asymptotic approximation to the distribution of the series-based partially linear model (PLM)
    - Accounts for the potentially large number of series-based terms→ more precise approximation
2. Identifies a problem with the variance of this estimator under "classical" asymptotics
    - Importantly, this problem also permeates to OLS with many covariates!!
3. Proposes solutions to the problem, allowing improved inference

New Asymptotic Approximation

- I am an applied person… do I need to know this…?

New Asymptotic Approximation

- I am an applied person… do I need to know this…?

  YES!!

New Asymptotic Approximation

- I am an applied person… do I need to know this…?

  YES!!

  … at least the general ideas (you can spare the details)

New Asymptotic Approximation

- I am an applied person… do I need to know this…?

  YES!!

  … at least the general ideas (you can spare the details)

- Usual asymptotics is “first-order”→ other expressions (including those depending on the # of terms/regressors) are assumed to vanish asymptotically (i.e., are ignored)

New Asymptotic Approximation

- I am an applied person… do I need to know this…?

  YES!!

  … at least the general ideas (you can spare the details)

- Usual asymptotics is "first-order"→ other expressions (including those depending on the # of terms/regressors) are assumed to vanish asymptotically (i.e., are ignored)

  - **Tension**:

  Need to assume K is large (in nonparametric estimation)… but "not too large" relative to the sample size such that $K/n \to 0$

New Asymptotic Approximation

- Cattaneo et al. perform a refinement of these asymptotics by accounting for the number of terms/regressors

New Asymptotic Approximation

- Cattaneo et al. perform a refinement of these asymptotics by accounting for the number of terms/regressors
    - Let the number of terms/regressors grow proportionally with the sample size (K/n converges to a constant)

New Asymptotic Approximation

- Cattaneo et al. perform a refinement of these asymptotics by accounting for the number of terms/regressors
  - Let the number of terms/regressors grow proportionally with the sample size (K/n converges to a constant)
  - Reduces the “tension” above and weakens the conditions

New Asymptotic Approximation

- Cattaneo et al. perform a refinement of these asymptotics by accounting for the number of terms/regressors
  - Let the number of terms/regressors grow proportionally with the sample size (K/n converges to a constant)
  - Reduces the “tension” above and weakens the conditions

- By working out the approximation and analyzing it:

New Asymptotic Approximation

- Cattaneo et al. perform a refinement of these asymptotics by accounting for the number of terms/regressors
  - Let the number of terms/regressors grow proportionally with the sample size (K/n converges to a constant)
  - Reduces the “tension” above and weakens the conditions
- By working out the approximation and analyzing it:
  - they obtain novel insights, and

New Asymptotic Approximation

- Cattaneo et al. perform a refinement of these asymptotics by accounting for the number of terms/regressors
  - Let the number of terms/regressors grow proportionally with the sample size (K/n converges to a constant)
  - Reduces the “tension” above and weakens the conditions

- By working out the approximation and analyzing it:
  - they obtain novel insights, and
  - a better approximation to the actual finite sample behavior!

Insights abut the Variance

- In practice:

  Using the series-based PLM with “classical” robust std. errors results in over-rejection of the null of no effect

Insights abut the Variance

- In practice:

  Using the series-based PLM with “classical” robust std. errors results in over-rejection of the null of no effect

- This **also** holds true in the OLS model (with robust std. errors) when many regressors are employed!!

Insights abut the Variance

- In practice:

  Using the series-based PLM with "classical" robust std. errors results in over-rejection of the null of no effect

- This **also** holds true in the OLS model (with robust std. errors) when many regressors are employed!!

  - Nowadays, using many covariates is typically the rule, rather than the exception

Insights abut the Variance

- In practice:

  Using the series-based PLM with "classical" robust std. errors results in over-rejection of the null of no effect

- This **also** holds true in the OLS model (with robust std. errors) when many regressors are employed!!
  - Nowadays, using many covariates is typically the rule, rather than the exception
    - Eg. when trying to control for "all relevant factors"

Insights abut the Variance

- In practice:

  Using the series-based PLM with "classical" robust std. errors results in over-rejection of the null of no effect

- This **also** holds true in the OLS model (with robust std. errors) when many regressors are employed!!
  - Nowadays, using many covariates is typically the rule, rather than the exception
    - Eg. when trying to control for "all relevant factors"
    - Wide availability of surveys with *many* variables

Insights abut the Variance

- In practice:

  Using the series-based PLM with “classical” robust std. errors results in over-rejection of the null of no effect

- This **also** holds true in the OLS model (with robust std. errors) when many regressors are employed!!
  - Nowadays, using many covariates is typically the rule, rather than the exception
    - Eg. when trying to control for “all relevant factors”
    - Wide availability of surveys with *many* variables
  - Interestingly, the “small sample” OLS std. errors under (true) homoskedasticity are correct

Insights abut the Variance

- In practice:

  Using the series-based PLM with "classical" robust std. errors results in over-rejection of the null of no effect

- This **also** holds true in the OLS model (with robust std. errors) when many regressors are employed!!
  - Nowadays, using many covariates is typically the rule, rather than the exception
    - Eg. when trying to control for "all relevant factors"
    - Wide availability of surveys with *many* variables
  - Interestingly, the "small sample" OLS std. errors under (true) homoskedasticity are correct
    - So, use this one with OLS when the number of regressors is large…

Solutions

- The authors propose two different solutions to estimate the variance in this context:

Solutions

- The authors propose two different solutions to estimate the variance in this context:
  - Both are robust to heteroskedasticity and many terms

Solutions

- The authors propose two different solutions to estimate the variance in this context:
  - Both are robust to heteroskedasticity and many terms
  - Account for the extra noise introduced by letting K grow with the sample size

Solutions

- The authors propose two different solutions to estimate the variance in this context:
  - Both are robust to heteroskedasticity and many terms
  - Account for the extra noise introduced by letting K grow with the sample size
  - Both "fix" the number of terms (K) in the given sample (e.g., selecting K using cross-validation)

Solutions

- The authors propose two different solutions to estimate the variance in this context:
  - Both are robust to heteroskedasticity and many terms
  - Account for the extra noise introduced by letting K grow with the sample size
  - Both "fix" the number of terms (K) in the given sample (e.g., selecting K using cross-validation)
- The authors show that these two solutions perform well in (preliminary) simulations, relative to the use of classical robust std. errors

Comments

- Can we say anything about *bootstrap* standard errors?

Comments

- Can we say anything about *bootstrap* standard errors?
- Would be great to implement the procedures in STATA

Comments

- Can we say anything about *bootstrap* standard errors?

- Would be great to implement the procedures in STATA

- Any way to sidestep the need for cross-validation in the choice of K?

Comments

- Can we say anything about *bootstrap* standard errors?
- Would be great to implement the procedures in STATA
- Any way to sidestep the need for cross-validation in the choice of K?
- An empirical implementation with a well known data set will also shed light on the results and implications

Comments

- Can we say anything about *bootstrap* standard errors?
- Would be great to implement the procedures in STATA
- Any way to sidestep the need for cross-validation in the choice of K?
- An empirical implementation with a well known data set will also shed light on the results and implications
- Another solution on the works…
  - Does it work better in simulations? What is the proportional improvement?

Final Thoughts

- The very nice thing about this paper…

Final Thoughts

- The very nice thing about this paper…
  - Moves us a step closer to robust models that do not sacrifice much of the attractive features of the linear model…

Final Thoughts

- The very nice thing about this paper…
  - Moves us a step closer to robust models that do not sacrifice much of the attractive features of the linear model…
  - …by making semipar. inference less dependent on arbitrary choices (e.g., choice of terms, bandwidth selection)

Final Thoughts

- The very nice thing about this paper…
  - Moves us a step closer to robust models that do not sacrifice much of the attractive features of the linear model…
  - …by making semipar. inference less dependent on arbitrary choices (e.g., choice of terms, bandwidth selection)
- There is still work to do…

Final Thoughts

- The very nice thing about this paper…
  - Moves us a step closer to robust models that do not sacrifice much of the attractive features of the linear model…
  - …by making semipar. inference less dependent on arbitrary choices (e.g., choice of terms, bandwidth selection)
- There is still work to do…
  - Tricks to make bandwidth/term selection less computationally intensive

Final Thoughts

- The very nice thing about this paper…
  - Moves us a step closer to robust models that do not sacrifice much of the attractive features of the linear model…
  - …by making semipar. inference less dependent on arbitrary choices (e.g., choice of terms, bandwidth selection)
- There is still work to do…
  - Tricks to make bandwidth/term selection less computationally intensive
  - Optimality criterion to select the number of terms in the series (work being conducted by Matías?)

Final Thoughts

- The very nice thing about this paper…
  - Moves us a step closer to robust models that do not sacrifice much of the attractive features of the linear model…
  - …by making semipar. inference less dependent on arbitrary choices (e.g., choice of terms, bandwidth selection)
- There is still work to do…
  - Tricks to make bandwidth/term selection less computationally intensive
  - Optimality criterion to select the number of terms in the series (work being conducted by Matías?)
  - Etc.

Final Thoughts

- The very nice thing about this paper…
  - Moves us a step closer to robust models that do not sacrifice much of the attractive features of the linear model…
  - …by making semipar. inference less dependent on arbitrary choices (e.g., choice of terms, bandwidth selection)
- There is still work to do…
  - Tricks to make bandwidth/term selection less computationally intensive
  - Optimality criterion to select the number of terms in the series (work being conducted by Matías?)
  - Etc.
- As of now: lots of room for applied econometricians to be creative in the implementation of these procedures
  - The "art" of applied econometrics !!