Bias Corrections in Testing and Estimating Semiparametric, Single Index Models

Roger Klein and Chan Shen

Rutgers University

2008 North American Summer Meeting of the Econometric Society
June 19th, 2008
Objective (Using Regular Kernels):

- Test: The Single Index Assumption
- A Bias Corrected SLS estimator
- Properties: Test & Estimator

Orthogonality Test: $\varepsilon \equiv Y - E(Y|v)$

$$E[\varepsilon F(X)] \not= 0$$

- Motivation: Single Index

$$E(Y|X) = E[Y|v], \ v = X\beta_0$$

- Problem: Selecting $F(X)$. 

Roger Klein and Chan Shen (RU)  Bias Corrections in Testing and Estimating Single Index Models 6/19/2008 2 / 19
Outline (cont.)

- Bias Control
  - Test: Recentering
  - Estimator: Two-Stage SLS

- Monte-Carlo Study
  - Estimator
  - Test Statistic
Test Statistic: $\hat{T} \equiv (Y - \hat{E}(Y|\hat{v}))\hat{F}(X_k)$

- **Assumptions**
  - i.i.d. observations
  - $X_1$ continuous
  - Null Model: Single Index

- **Estimated Index:** $\hat{v} = X_1 + X_2\hat{\theta}$

- **$\hat{F}(X_k) \equiv \hat{r}\hat{E}(Y|X_k)$**
  - Avoiding Degeneracy ($v$ or $x$)
  - Dimensionality ($k$)
  - General Linear Model
Bias Reduction in Testing

\[ \hat{T}^* = (Y - \hat{E}(Y|\hat{v})) (\hat{F}(x_k) - \hat{E}[\hat{F}(x_k)|\hat{v}]) \]

- Motivation: Recentering for Bias Control

- Large-Sample Equivalence in Test Statistics:(U-Statistics)

\[ \sqrt{N} [\hat{T}^* - \hat{T}] \to 0 \]

- Monte-Carlo: Recentering Matters.
The Estimator

- SLS (Ichimura):
  \[ \hat{\theta} = \arg \min \sum_i \hat{\tau}(x_i) \left[ Y_i - \hat{E}[Y_i|X_1i + X_2i\theta] \right]^2 \]

\[ \sqrt{N} \hat{G}(\theta_0) = A : \frac{1}{\sqrt{N}} \sum_i \tau(x_i) \left[ Y_i - E_i \right] \nabla E_i \]

\[ B : -\frac{1}{\sqrt{N}} \sum_i \left[ \hat{E}_i - E_i \right] \hat{\tau}(x_i) \nabla \hat{E}_i + o_p(1) \]

- A: Required form.

- B: Bias Problem
  - Higher order Kernels \([\hat{E}_i - E_i]\)
  - "Residual": \(E[\nabla E|Index] = 0\), exploitable with \(\tau(\hat{v})\)
  - A smoothing adjustment \((\hat{E}_i^*)\)
Bias Reduction in SLS Estimation: A 2-Stage Strategy

- **Stage 1: SLS Estimation, \( \hat{\tau}(x_i) \)**
  - Consistency: \( \hat{E} = \hat{f} / \hat{g} \)
  - Below \( \sqrt{N} \)-rate convergence

- **Stage 2: SLS Estimation, \( \hat{\tau}(\hat{v}_i) \)**
  - Consistency: \( \hat{E_a} = \hat{f} / [\hat{g} + \Delta_N] \)
  - \( \sqrt{N} \)-rate convergence: Residual Bias Control
A Smoothing Adjustment

- Modified Gradient Component

\[-\frac{1}{N} \sum_i [\hat{E}_i - E_i] \tau(v_i) \nabla \hat{E}_i, \ r < 1/6\]

\[-\frac{1}{N} \sum_i [\hat{E}_i^* - E_i] \tau(v_i) \nabla \hat{E}_i, \ r^* = 1/5\]

\[\hat{\theta}^* = \hat{\theta} - \hat{H}^{-1} \hat{A}\]

\[\hat{A} = \left[ \frac{1}{N} \sum_i [\hat{E}_i - \hat{E}_i^*] \tau(v_i) \nabla \hat{E}_i \right]\]
Monte Carlo Designs: $N, \text{Rep} = 1000$

Three Parameter Design

- $H_0:$

  $$Y_i = M_{0i} + \varepsilon_i, \quad M_{0i} \propto (X_{1i} + X_{2i} + 2X_{3i})^2$$

  where the $X_i$s $\sim \chi^2(1)$ and $\varepsilon \sim N(0,1)$.

- $H_1$ (No Index Structure):

  $$Y_i = M_{1i} + \varepsilon_i, \quad M_{1i} \propto [3X_{1i}^2 + 2X_{2i}^2 + X_{3i}^2 + 3]$$
Basic Design

- $H_0$: 
  \[ Y_i = M_{0i} + \varepsilon_i, \quad M_{0i} \propto (X_{1i} + X_{2i})^2 \]
  where the $X_i$s $\sim \chi^2(1)$ and $\varepsilon \sim N(0, 1)$.

- $H_1$: 
  \[ Y_i = M_{1i} + \varepsilon_i, \quad M_{1i} \propto \left[ (X_{1i} + X_{2i})^2 + (X_{1i} + X_{2i})^3 \right] \]
Binary Response Design

- **H₀:**
  \[ Y_i = \begin{cases} 
  1 : & M_{0i} > \epsilon_i, \quad M_{0i} \propto X_{1i} + X_{2i} - 0.5 \\
  0 : & \text{otherwise}
  \end{cases} \]

  - The two X’s are correlated. In particular, the X’s are both linear combinations of the same \( \chi^2 \) variable and different normal shocks.

- **H₁:**
  \[ Y_i = \begin{cases} 
  1 : & M_{0i} > M_{1i} \epsilon_i, \quad M_{1i} \propto \sqrt{1 + (X_{1i} - X_{2i})^2} \\
  0 : & \text{otherwise}
  \end{cases} \]
Discrete Regressor Design: $X_2$ Binary

- $H_0$:
  \[ Y_i = M_{0i} + \varepsilon_i, \quad M_{0i} \propto (X_{1i} + X_{2i})^2. \]

- $H_1$:
  \[ Y_i = M_{1i} + \varepsilon_i, \quad M_{1i} \propto [(X_{1i} + X_{2i})^2 + (X_{1i} + X_{2i})^3] \]
Estimation Results

<table>
<thead>
<tr>
<th>Three Parameter Design*</th>
<th>S1SLS</th>
<th>S2SLS</th>
<th>CS2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td>-0.005</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Rvar</td>
<td>0.043</td>
<td>-0.003</td>
<td>-0.004</td>
</tr>
<tr>
<td>Rmse</td>
<td>0.089</td>
<td>0.069</td>
<td>0.069</td>
</tr>
</tbody>
</table>

* Twicing kernel results removed.
## Discrete Regressor Design

<table>
<thead>
<tr>
<th></th>
<th>SLS-TW</th>
<th>S1SLS</th>
<th>S2SLS</th>
<th>CS2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td>-0.019</td>
<td>-0.023</td>
<td>-0.036</td>
<td>-0.019</td>
</tr>
<tr>
<td>Rvar</td>
<td>0.045</td>
<td>0.041</td>
<td>0.036</td>
<td>0.037</td>
</tr>
<tr>
<td>Rmse</td>
<td>0.049</td>
<td>0.047</td>
<td>0.050</td>
<td>0.042</td>
</tr>
</tbody>
</table>
## Test Results

### Three Parameter Design

<table>
<thead>
<tr>
<th></th>
<th>Uncentered</th>
<th>Recentered</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW size</td>
<td>0.132</td>
<td>0.045</td>
</tr>
<tr>
<td>TW power</td>
<td>0.822</td>
<td>0.806</td>
</tr>
<tr>
<td>TW adjusted power</td>
<td>0.697</td>
<td>0.817</td>
</tr>
<tr>
<td>BRR size</td>
<td>0.153</td>
<td>0.049</td>
</tr>
<tr>
<td>BRR power</td>
<td>0.909</td>
<td>0.817</td>
</tr>
<tr>
<td>BRR adjusted power</td>
<td>0.806</td>
<td>0.819</td>
</tr>
</tbody>
</table>
## Basic Design

<table>
<thead>
<tr>
<th></th>
<th>Uncentered</th>
<th>Recentered</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TW</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>0.232</td>
<td>0.063</td>
</tr>
<tr>
<td>power</td>
<td>0.899</td>
<td>0.890</td>
</tr>
<tr>
<td>adjusted power</td>
<td>0.665</td>
<td>0.878</td>
</tr>
<tr>
<td><strong>BRR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>0.232</td>
<td>0.065</td>
</tr>
<tr>
<td>power</td>
<td>0.905</td>
<td>0.889</td>
</tr>
<tr>
<td>adjusted power</td>
<td>0.744</td>
<td>0.878</td>
</tr>
<tr>
<td></td>
<td>Uncentered</td>
<td>Recentered</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>TW</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>0.125</td>
<td>0.022</td>
</tr>
<tr>
<td>power</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>adjusted power</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td><strong>BRR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>0.167</td>
<td>0.023</td>
</tr>
<tr>
<td>power</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>adjusted power</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
### Discrete Regressor Design

<table>
<thead>
<tr>
<th></th>
<th>Uncentered</th>
<th>Recentered</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TW</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>0.770</td>
<td>0.177</td>
</tr>
<tr>
<td>power</td>
<td>0.996</td>
<td>0.994</td>
</tr>
<tr>
<td>adjusted power</td>
<td>0.069</td>
<td>0.986</td>
</tr>
<tr>
<td><strong>BRR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>0.874</td>
<td>0.053</td>
</tr>
<tr>
<td>power</td>
<td>0.999</td>
<td>0.997</td>
</tr>
<tr>
<td>adjusted power</td>
<td>0.144</td>
<td>0.997</td>
</tr>
</tbody>
</table>
Conclusions and Directions for Future Research

- Large Sample Properties
Conclusions and Directions for Future Research

- Large Sample Properties
  - Test statistic (quadratic form) follows $\chi^2$
Conclusions and Directions for Future Research

- Large Sample Properties
  - Test statistic (quadratic form) follows $\chi^2$
  - Estimator achieves consistency, linearity and normality
Conclusions and Directions for Future Research

- Large Sample Properties
  - Test statistic (quadratic form) follows $\chi^2$
  - Estimator achieves consistency, linearity and normality

- Finite Sample Properties (Monte Carlo)
Conclusions and Directions for Future Research

- Large Sample Properties
  - Test statistic (quadratic form) follows $\chi^2$
  - Estimator achieves consistency, linearity and normality

- Finite Sample Properties (Monte Carlo)
  - Test statistic has good size and power properties
Conclusions and Directions for Future Research

- Large Sample Properties
  - Test statistic (quadratic form) follows $\chi^2$
  - Estimator achieves consistency, linearity and normality

- Finite Sample Properties (Monte Carlo)
  - Test statistic has good size and power properties
  - Estimator dominates the others