Institutions and Growth: Evidence from Estimation Methods Robust to Weak Instruments

2009 North American Summer Meeting of the Econometric Society

Alexandre Dmitriev
UNSW

Boston, June 2009
Institutions and Growth: "Why do some countries produce so much more output per worker than others?"

Hall and Jones (QJE 1999) Linear IV Regression Model

$$\log(Y/L) = \alpha + \beta S + \varepsilon$$

$$S = Z\delta + u$$

Proxy for social infrastructure ($S$): GADP + openness measure.

Instruments ($Z$):

- The fraction of population speaking English at birth;
- The fraction of population speaking one of the five major European languages at birth;
- The distance from the equator;
- The Frankel and Romer (1999) geography predicted trade intensity.
Acemoglu et al (AER 2001) direct criticism towards HJ99 for reliance on the two instruments with questionable relation to the proposed theory: distance from the equator and Frankel-Romer predicted trade share.

"Hall and Jones (1999), in turn, use distance from the equator as an instrument for social infrastructure because, they argue, latitude is correlated with Western influence, which leads to good institutions. The theoretical reasoning for these instruments is not entirely convincing.” Acemoglu et al (2001:1373)
Critique
The Instruments

- Frankel and Romer (1999) predicted trade intensity: no justification
- "latitude": Hall and Jones (1999:101) argue that Europeans were "...more likely to settle in areas that were broadly similar in climate to West Europe"
- Linguistic instruments: The extent to which primary languages of Western Europe are spoken as first language today reflects the extent of positive Western European influence during 16-19th centuries
A problem:

- The instruments proposed by Hall and Jones (1999) and in particular linguistic instruments are only weakly correlated with their proxy for institutional quality.

Two Questions:

- Can European languages spoken at birth instrument for institutional quality?
- Are the results of Hall and Jones (1999) driven by their reliance on "latitude" as an instrument?
Weak Instruments

- Under weak instruments the sampling distributions of IV statistics are non-normal;
- IV point estimates, hypothesis tests and confidence intervals are unreliable;
- Weak instruments is not a small sample problem; (Bound, Jager and Baker, 1995)
- Basic Diagnostics: First-Stage $F$-statistics; (Staiger and Stock, 1997)
### Basic Diagnostic for Weak Instruments

**Table:** First-stage $F$-statistics across specifications

<table>
<thead>
<tr>
<th>Specification</th>
<th>First-stage $F$ statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) Full sample (127 countries), 4 instruments</td>
<td>12.843</td>
</tr>
<tr>
<td>ii) Full sample (127 countries), latitude &amp; FR tr.share</td>
<td>13.118</td>
</tr>
<tr>
<td>iii) Full sample (127 countries), languages</td>
<td>6.4995</td>
</tr>
<tr>
<td>iv) No imputed data (79 countries), 4 instruments</td>
<td>6.7206</td>
</tr>
<tr>
<td>v) No imputed data (79 countries), latitude &amp; FR tr.share</td>
<td>9.0664</td>
</tr>
<tr>
<td>vi) No imputed data (79 countries), languages</td>
<td>2.1178</td>
</tr>
</tbody>
</table>
Even if $F$-statistics are low (less than 5) the instruments do not have to be irrelevant (Staiger and Stock, 1997).

Some methods are partially robust to weak instruments in the sense that they are more reliable than TSLS when instruments are weak; (Stock and Yogo, 2003)

The ranking of the partially robust estimators heavily depends on the nature of the data. (Blomquist and Dahlberg, 1999)
Monte Carlo Experiments

The linear IV regression model estimated by Hall and Jones (1999)

\[
\log(Y/L) = \alpha + \beta S + \varepsilon \\
S = Z\delta + u
\]

The dependent variable is generated as

\[
y = a + bS + e
\]

where \( a = 7 \), \( b = 3 \), and \( S \) is the proxy for social infrastructure from HJ99 study.

The error term is generated by \( e = k\hat{u} + \xi \) where \( \xi \sim N(0, \sigma^2_{\xi}) \) and \( \hat{u} \) is a projection of \( S \) to the space orthogonal to the space spanned by the instruments, i.e. \( \hat{u} = (I - Z(Z'Z)^{-1}Z')S \).

The values of the parameters \( \sigma^2_{\xi} = 0.1 \), and \( k = 0.8 \) are set so that first two empirical moments of generated \( y \) would be close to those of the observed \( \log(Y/L) \).
Monte Carlo Experiments

- Performance of TSLS estimator does not seem to be severely affected by the weak correlation of the instruments with the endogenous regressor.

- Jackknife estimators are outranked by the k-class estimators. In particular JIVE suffers from both severe average bias and size distortion.

- Shifting from the four instruments to the linguistic instruments only makes the estimates much more imprecise.

- Specification vi):
  - performance of OLS, LIML and Jackknife estimators is inferior both in terms of bias, RMSE, and median absolute error.
  - Fuller - k outranks the remaining estimators in terms of median square error, RMSE, the size distortion.
Monte Carlo Experiments
Larger Sample, All Four Instruments

- Sample: 129 Countries, Some imputed data
- Instruments: distance to the equator; geography predicted trade intensity; two linguistic instruments
Monte Carlo Experiments
Smaller Sample, Two Linguistic Instruments Only

- Sample: 79 Countries, No imputed data
- Instruments: Fraction of population speaking English at birth; Fraction of population speaking one of five European languages at birth;
Monte Carlo Experiments
Smaller Sample, Two Linguistic Instruments Only

<table>
<thead>
<tr>
<th>$k$–class estimators</th>
<th>OLS</th>
<th>TSLS</th>
<th>LIML</th>
<th>Fuller–$k$</th>
<th>JIVE</th>
<th>UJIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification vi)</td>
<td>(No imputed data, languages); $\rho = 0.516$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Av. % Bias</td>
<td>0.251</td>
<td>-0.010</td>
<td>-0.055</td>
<td>0.017</td>
<td>-0.887</td>
<td>-0.616</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.766</td>
<td>0.599</td>
<td>0.912</td>
<td>0.535</td>
<td>2.724</td>
<td>2.699</td>
</tr>
<tr>
<td>SIZE</td>
<td>0.000</td>
<td>0.998</td>
<td>1.000</td>
<td>0.989</td>
<td>0.543</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: Average per cent bias (Av % Bias), RMSE and SIZE (coverage rate for a 95% confidence interval calculated as a proportion of replications when the confidence interval covers the true parameter value) for the estimates of $\beta$. Reported $\rho$ indicates the average sample correlation coefficient between the endogenous regressor and the error term. The experiments rely on 1000 Monte Carlo replications and 1000 Bootstrap iterations.
## Estimation Results: Partially Robust Estimators

<table>
<thead>
<tr>
<th></th>
<th>$k-$class estimators</th>
<th>Jackknife estimators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>TSLS</td>
</tr>
<tr>
<td>i) 127 obs.; 4 inst</td>
<td>3.2891 (0.196)</td>
<td>5.0847 (0.5079)</td>
</tr>
<tr>
<td>ii) 127; Dist/FR</td>
<td>4.6698 (0.6763)</td>
<td>4.7600 (1.4348)</td>
</tr>
<tr>
<td>iii) 127; Languages</td>
<td>5.7981 (1.6004)</td>
<td>6.2097 (3.2017)</td>
</tr>
<tr>
<td>iv) 79; 4 inst.</td>
<td>3.0741 (0.6121)</td>
<td>4.6612 (3.7255)</td>
</tr>
<tr>
<td>v) 79; Dist./FR</td>
<td>3.9388 (0.7048)</td>
<td>3.9751 (7.9828)</td>
</tr>
<tr>
<td>vi) 79; Languages</td>
<td>6.538 (5.1294)</td>
<td>8.6454 (26.941)</td>
</tr>
</tbody>
</table>
Hall and Jones (1999) results are not driven by reliance on the "geographical" instruments.

In the case of the HJ99 study the TSLS and Fuller-$k$ estimators are plagued by neither severe size distortion or bias despite low values of first-stage $F$-statistics.

Some $k$-class estimators allow us to utilize the linguistic variables to instrument for institutional quality despite their low correlation with the endogenous regressor.