A LM test for the mean stationarity assumption in dynamic panel data models
The xttestms command

Laura Magazzini
Institute of Economics and EMbeDS, Sant’Anna School of Advanced Studies

(joint work with G. Calzolari, University of Firenze)

STATA Conference 2021, August 5
Outline

- Introduction
- GMM estimation of dynamic panel data models
- LM test for verifying initial conditions
- Monte Carlo simulation
- The `xttestms` command
- Examples
- Discussion
Dynamic panel data framework ($i = 1, ..., N$, $t = 1, ..., T$):

$$y_{it} = \rho y_{it-1} + x_{it}' \beta + \tau_t + u_i + e_{it}$$

Estimation relies on GMM methods to tackle the endogeneity of $y_{it-1}$

- Strictly exogenous, predetermined, simultaneous $x_{it}$

Identifying assumption: $e_{it}$ is uncorrelated over time

- Arellano & Bond (1991) test for residuals autocorrelation

Difference GMM estimator (AB91); non-linear estimator by Ahn & Schmidt (1995)

Blundell & Bond (1998) adds an assumption on initial conditions: system GMM estimator
GMM estimation

To simplify, $y_{it} = \rho y_{it-1} + u_i + e_{it} = \rho y_{it-1} + \varepsilon_{it}$

- To apply GMM, take first difference to remove $u_i$

  $$\Delta y_{it} = \rho \Delta y_{it-1} + \Delta e_{it}$$

- Difference GMM estimator (AB91): under the lack of autocorrelation in $e_{it}$ lag 2 or more of $y$ can be used as instrument for $\Delta y_{it}$

  $$E(y_{it-j} \Delta e_{it}) = 0 \ (t = 2, \ldots, T; \ j \geq 2)$$

- The non-linear GMM (AS95) estimator also considers

  $$E(\Delta \varepsilon_{it} \varepsilon_{iT}) = 0 \ \text{for every} \ t < T$$

- Efficiently exploits all available moment conditions
- So far, limited application in empirical analysis
GMM estimation

To simplify, \( y_{it} = \rho y_{it-1} + u_i + e_{it} = \rho y_{it-1} + \varepsilon_{it} \)

- SYS GMM (BB98) further exploits moment conditions on the “level”
equations:

\[
E(\Delta y_{it-1}\varepsilon_{it}) = 0
\]

- Effectively a condition on the initial observation (Roodman, 2009)
- If satisfied, outperform DIF GMM, especially with persistent processes
  (i.e. \( \rho \) close to 1 or \( \sigma_u^2 \) “large” w.r.t. \( \sigma_e^2 \))
- Validity of these additional moment conditions is usually tested on the
  basis of the difference between SYS GMM and DIF GMM
- Magazzini & Calzolari (2020) propose a different framework with
  better power in detecting violation of this assumption
The LM test for testing initial conditions
(Magazzini & Calzolari, 2020)

- The LM test treats the system GMM estimator as the restricted estimator in an “augmented” set of moment conditions
- If the “level” moment conditions are not satisfied, we can write:

\[ E(\Delta y_{it-1}\varepsilon_{it}) - \psi_{t-1} = 0 \]

- SYS GMM under \( H_0 : \psi_1 = \psi_2 = ... = \psi_{T-1} = 0 \)
- Asy. equivalent to diff-in-Hansen test comparing SYS and DIFF GMM

- MC20 notice that

\[ E(\Delta y_{it-1}\varepsilon_{it}) - \psi_{t-1} = E(\Delta y_{it-1}\varepsilon_{it}) - \rho^{t-2} \psi_1 = 0 \]

- In the pure dynamic framework, asy. equivalent to diff-in-Hansen test comparing SYS and NL GMM
- Larger power with respect to the customarily applied procedures (↓ dof)
The LM test for testing initial conditions

\[ y_{it} = \rho y_{it-1} + x_{it}' \beta + \varepsilon_{it} \]

- In the more general case

\[ \Delta y_{it} = \rho^{t-1} \Delta y_{i1} + \sum_{s=0}^{t-2} \rho^s (\Delta x_{it-s} \beta + \Delta \varepsilon_{it-s}) \]

- Strictly exogenous regressors: no additional moment condition from the level equations

- With predetermined or simultaneously determined \( x_{it} \), additional parameters should also be considered for the moment conditions related to \( x_{it} \)
  
  ▶ For example, in the case of a predetermined regressor, \( x_{it} \):

  \[ E(\Delta x_{it} \varepsilon_{it}) - \xi_t = 0 \]

  ▶ SYS GMM when \( \psi_1 = \xi_2 = \ldots = \xi_T = 0 \)
The LM test for testing initial conditions

\[ y_{it} = \rho y_{it-1} + x_{it}' \beta + \tau_t + \varepsilon_{it} \]

- The SYS GMM obtained as a restricted estimate in a set of “augmented” moment conditions (MC20)
- An LM strategy can be applied, computed on the basis of the SYS GMM estimates
- Computation of the LM test is based on the value of the gradient for the unconstrained criterion function evaluated at the restricted estimator (Newey & West, 1987; Ruud, 2000)

\[
LM = Ng_N(\hat{\theta}_{RN})'\hat{\Omega}^{-1}\hat{G}_N \left( \hat{G}'_N\hat{\Omega}^{-1}\hat{G}_N \right)^{-1} \hat{G}'_N\hat{\Omega}^{-1}g_N(\hat{\theta}_{RN})
\]

- \( \theta_{RN} \) includes \( \rho, \beta \) and the additional parameter (set to 0 under \( H_0 \))
- \( G_N = \partial g_N/\partial \theta \) has to be “augmented” with the additional parameters
- \( \Omega^{-1} \) corresponds to the weighting matrix of the SYS GMM
Monte Carlo set up

- $y_{it} = \rho y_{it-1} + x'_{it}\beta + \varepsilon_{it} = \rho y_{it-1} + x'_{it}\beta + u_i + e_{it}$
  - $u_i \sim N(0, \sigma^2_u)$
  - $e_{it} = \delta_i \tau_t \nu_{it}$ with $\delta_i \sim U(0.5, 1.5)$, $\tau_t \sim 0.5 + 0.1 t$, and $\nu_{it} \sim \chi^2_1 - 1$ (W05)
- The regressor $x_{it} = \rho_x x_{it-1} + \theta_u u_i + \theta_e \nu_{it} + w_{it}$
  - $\theta_u = 0.25$, $\theta_e = -0.1$, $w_{it} \sim N(0, 0.16)$ (BBW01)
  - We set $\rho = \rho_x = 0.5$ and $\beta = 1$
  - $x_{it}$ as strictly exogenous ($\nu_{it} \sim N(0, 1)$) or simultaneously determined ($\nu_{it} = e_{it}$)
- Departure from mean stationarity by the parameters $\gamma_y$ and $\gamma_x$ that multiply the individual component in the initial observations
  - Condition on initial observation satisfied if $\gamma_y = \gamma_x = 1$
Monte Carlo results - *xtdpdsys*

Strictly exogenous $x_{it}$

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>$\gamma_x$</th>
<th>$\gamma_y$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{\beta}$</th>
<th>$H$</th>
<th>diffH</th>
<th>LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>4</td>
<td>1.0</td>
<td>1.0</td>
<td>0.515</td>
<td>0.999</td>
<td>3.40</td>
<td>5.22</td>
<td>4.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.139)</td>
<td>(0.248)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>8</td>
<td>1.0</td>
<td>1.0</td>
<td>0.507</td>
<td>1.004</td>
<td>1.76</td>
<td>5.65</td>
<td>5.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.064)</td>
<td>(0.172)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td>0.6</td>
<td>0.6</td>
<td>0.732</td>
<td>1.068</td>
<td>9.97</td>
<td>16.95</td>
<td>19.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.083)</td>
<td>(0.262)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>8</td>
<td>0.6</td>
<td>0.6</td>
<td>0.612</td>
<td>1.127</td>
<td>18.24</td>
<td>52.55</td>
<td>74.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.062)</td>
<td>(0.180)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td>1.4</td>
<td>1.4</td>
<td>0.713</td>
<td>1.150</td>
<td>41.34</td>
<td>55.42</td>
<td>62.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.118)</td>
<td>(0.276)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>8</td>
<td>1.4</td>
<td>1.4</td>
<td>0.535</td>
<td>1.157</td>
<td>44.44</td>
<td>78.83</td>
<td>94.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.064)</td>
<td>(0.184)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- With $T = 4$, 7 m.c.; dof: $H = 4$, diffH = 2, LM = 1
- With $T = 8$, 29 m.c.; dof: $H = 26$, diffH = 6, LM = 1
Monte Carlo results - *xtdpdsys*

Simultaneously determined $x_{it}$

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>$\gamma_x$</th>
<th>$\gamma_y$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{\beta}$</th>
<th>H</th>
<th>diffH</th>
<th>LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>4</td>
<td>1.0</td>
<td>1.0</td>
<td>0.509</td>
<td>1.017</td>
<td>4.66</td>
<td>5.61</td>
<td>5.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.065)</td>
<td>(0.303)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>8</td>
<td>1.0</td>
<td>1.0</td>
<td>0.506</td>
<td>0.960</td>
<td>3.78</td>
<td>6.69</td>
<td>6.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.029)</td>
<td>(0.157)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>4</td>
<td>0.6</td>
<td>0.6</td>
<td>0.679</td>
<td>1.585</td>
<td>20.96</td>
<td>35.76</td>
<td>31.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.047)</td>
<td>(0.242)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>8</td>
<td>0.6</td>
<td>0.6</td>
<td>0.633</td>
<td>1.489</td>
<td>74.73</td>
<td>98.28</td>
<td>92.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.026)</td>
<td>(0.141)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>4</td>
<td>1.4</td>
<td>1.4</td>
<td>0.674</td>
<td>1.750</td>
<td>79.60</td>
<td>90.51</td>
<td>90.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.047)</td>
<td>(0.286)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>8</td>
<td>1.4</td>
<td>1.4</td>
<td>0.574</td>
<td>1.586</td>
<td>99.94</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.030)</td>
<td>(0.171)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- With $T = 4$, 11 m.c.; dof: $H = 8$, $\text{diffH} = 4$, $LM = 3$
- With $T = 8$, 55 m.c.; dof: $H = 52$, $\text{diffH} = 12$, $LM = 7$
The `xttestms` command

- After estimating the SYS GMM estimator using `xtdpdsys` or `xtabond2`, type:
  
  ```
  xttestms, [showgmm]
  ```

- Matrices to build the $LM$ statistics are obtained by `xtabond2`, ..., `svmat`
  - The model is re-estimated if necessary
  - If `showgmm` is specified, the re-estimated model is shown
Example 1

- Data used in Cameron and Trivedi (2005, ch. 21-22), taken from Ziliak (1997)
- Labour supply of 532 individuals over the years 1979-1988
- Dependent variable: $lnhrs$, the log of annual hours worked
- Regressor: $lnwg$, the natural log of hourly wage
  - Dynamic specification with no additional regressors
  - $lnwg$ as strictly exogenous, predetermined, simultaneously determined
Example 1: a labour equation

Dynamic model with no regressors: $lnhr_{it} = \mu + \rho lnhr_{it-1} + \tau_t + u_i + e_{it}$

- Estimate SYS GMM:
  xtdpdsys lnhr dyear3-dyear10, twostep vce(robust)
- AB91 test does not reject the null hypothesis of lack of autocorrelation in the residuals $e_{it}$
  $\nabla AR1 = -3.55 \ (p < 1\%)$ and $AR2 = 0.14 \ (p = 0.89)$
- After the estimation, the LM test can be computed by typing xttestms:

  Number of lags detected in the equation: 1
  lag(s) of lnhr included among the regressors: 1

  LM test of mean stationarity
  Test = 6.82063 with p-value .009011
  The test has a chi2(1) distribution
Example 1: dynamic model with no regressors

“Augmented” m.c.: \( E(\Delta y_{it-1} \varepsilon_{it}) - \rho^{t-2} \psi_1 = 0 \)

```
. mata: mata set matafavor speed
. xtabond2 lnhr l.lnhr dyear3-dyear10, gmmstyle(l.lnhr) h(2) ///
   ivstyle(dyear3-dyear10, eq(level)) twostep robust svmat
...
```

```
. mat G=-(e(Z))’*(e(X))
. mat gpsi = J(colsof(e(Z),1,0)
. mat gpsi[colnumb(e(Z), "Levels eq:LD.lnhr/1981"),1]=-_b[L.lnhr]^0
. mat gpsi[colnumb(e(Z), "Levels eq:LD.lnhr/1982"),1]=-_b[L.lnhr]^1
. mat gpsi[colnumb(e(Z), "Levels eq:LD.lnhr/1983"),1]=-_b[L.lnhr]^2
...
. mat gpsi[colnumb(e(Z), "Levels eq:LD.lnhr/1988"),1]=-_b[L.lnhr]^7
. mat G=(G,gpsi)
. mat testcm = e(Ze)’*e(A2)*G*invsym(G’*e(A2)*G)*G’*e(A2)*e(Ze)
```

- Hansen test of overid. restrictions, equal to 68.26 with \( p \)-value 0.008
- Difference-in-Hansen test, equal to 16.22 with \( p \)-value 0.039
Example 1
Including $lnwg$ in the equation

- Treat $lnwg$ as strictly exogenous
  
  . xtdpdsys lnhr lnwg dyear3-dyear10, twostep vce(robust)

  [output omitted]

  . xttestms
  Number of lags detected in the equation: 1
  lag(s) of lnhr included among the regressors: 1

  LM test of mean stationarity
  Test = 7.02113 with p-value  .008055
  The test has a chi2(1) distribution
Example 1
Including \textit{lnwg} in the equation

- Treat \textit{lnwg} as predetermined: \texttt{xtdpdsys lnhr dyear3-dyear10, twostep vce(robust) pre(lnwg)}
  
  \texttt{. xttestms}
  
  Number of lags detected in the equation: 1
  
  lag(s) of \texttt{lnhr} included among the regressors: 1
  
  lag(s) of \texttt{lnwg} included among the regressors: 0

  LM test of mean stationarity
  
  Test = 14.7368 with p-value .141955
  
  The test has a chi2(10) distribution
Example 1
Including \( lnwg \) in the equation

- The test has 10 degrees of freedom as we are also considering the “augmented” moment conditions related to \( x_{it} \)

\[
E(\Delta x_{it} \epsilon_{it}) - \xi_t = 0
\]

- By the recursive formula, these parameters also enter the m.c. related to \( y_{it-1} \)

\[
E(\Delta \ln hr_{i,t \epsilon i,80}) = \psi_1
\]
\[
E(\Delta \ln hr_{i,t \epsilon i,81}) = E[(\rho \Delta \ln hr_{80} + \beta \Delta \ln wg_{81} + \Delta e_{81}) \epsilon_{82}]
\]
\[
= \rho \psi_1 + \beta E(\Delta \ln wg_{81} \epsilon_{82}) = \rho \psi_1 + \beta \xi_2
\]
\[
\vdots
\]
\[
E(\Delta \ln hr_{i,t \epsilon i,87}) = \rho^7 \psi_1 + \beta (\rho^6 \xi_2 + \rho^5 \xi_3 + \ldots + \xi_8)
\]
Example 1
Including *lnwg* in the equation

- Treat *lnwg* as simultaneously determined:
  . *xtdpd*sys *lnhr* dyear3-dyear10, endog(*lnwg*) twostep
  vce(robust)

- After the estimation, the LM test for mean stationarity can be invoked by using *xttestms*:
  . *xttestms*
  Number of lags detected in the equation: 1
  lag(s) of *lnhr* included among the regressors: 1
  lag(s) of *lnwg* included among the regressors: 0

  LM test of mean stationarity
  Test = 6.70805 with p-value .667486
  The test has a chi2(9) distribution
Example 2

- usbal89.dta by Blundell & Bond (2000) and Bond (2002)
- Balanced panel dataset of 509 US firms observed over 8 years, 1982-1989
- The estimated equation is
  ```
  .* xi: xtabond2 y l.y n l.n k l.k i.year , ///
  gmm(y n k, lag(3 .)) iv(i.year, equation(level))
  twostep robust
  ▶ Only lags 3 or older can be used as legitimate instruments
  ▶ Lagged values of the regressors are included in the equation of interest
  ▶ Preferred specification: n and k as simultaneously determined
  ```
Example 2

Standard diagnostics & \texttt{xttestms}

Arellano-Bond test for AR(1) in first differences: $z = -7.90$ Pr $> z = 0.000$
Arellano-Bond test for AR(2) in first differences: $z = -0.58$ Pr $> z = 0.559$

Hansen test of overid. restrictions: $\text{chi2}(55) = 79.45$ Prob $> \text{chi2} = 0.017$

Difference-in-Hansen tests of exogeneity of instrument subsets:
\begin{itemize}
  \item GMM instruments for levels
    \begin{itemize}
      \item Hansen test excluding group: $\text{chi2}(40) = 38.33$ Prob $> \text{chi2} = 0.546$
      \item Difference (null H = exogenous): $\text{chi2}(15) = 41.12$ Prob $> \text{chi2} = 0.000$
    \end{itemize}
\end{itemize}

\texttt{. xttestms}

Number of lags detected in the equation: 1
\begin{itemize}
  \item lag(s) of y included among the regressors: 1
  \item lag(s) of n included among the regressors: 0 1
  \item lag(s) of k included among the regressors: 0 1
\end{itemize}

LM test of mean stationarity
\begin{itemize}
  \item Test $= 33.3191$ with p-value $= 0.000467$
  \item The test has a chi2(11) distribution
\end{itemize}
LM test to better assess validity of initial condition in SYS GMM
Outperform customarily employed testing procedures
  ▶ In the pure dynamic case, the proposed procedure contrasts SYS and NL GMM
  ▶ Better performance in the case of strictly exogenous regressors
  ▶ Further work should consider alternative routes to detecting departures from mean stationarity in the case of “endogenous” regressors
Main references


Main references


CT05 Cameron, A. C., and Trivedi, P. K.: 2005, Microeconometrics: Methods and Applications, Cambridge University Press.


Thank you

laura.magazzini@santannapisa.it