Robust testing for serial correlation in linear panel-data models

Sebastian Kripfganz

University of Exeter Business School, Department of Economics, Exeter, UK
www.kripfganz.de @Kripfganz

30th London Stata Conference

September 13, 2024

University of Exeter Business School

```
ssc install xtdpdserial
net install xtdpdserial, from(http://www.kripfganz.de/stata/)
```

Linear panel data model

- Consider the linear panel data model

$$y_{it} = \mathbf{x}_{it}'\beta + \underbrace{\alpha_i + u_{it}}_{=e_{it}}$$

  - The regressors $\mathbf{x}_{it}$ can be strictly exogenous, predetermined, or endogenous, and they might include a lagged dependent variable $y_{i,t-1}$.
- It is standard practice to test for serial correlation in the idiosyncratic error component u_{it}.
  - In the best case, serial correlation just leads to a loss of efficiency and requires standard errors to be made panel/cluster-robust (which should be done by default anyways).
  - In the worst case, serial correlation can cause inconsistency of the coefficient estimator $\hat{\beta}$, which is the typical case in dynamic models.

Existing serial correlation tests in Stata

- For static models with strictly exogenous regressors, several tests are available in Stata (some of them assuming homoskedasticity).
  - `xtserial` (Drukker, 2003): Wooldridge-Drukker test for $Corr(\Delta e_{it}, \Delta e_{i,t-1}) = -0.5$
  - `xtistest`, `xtqptest`, `xthrtest` (Wursten, 2018) after `xtreg`: tests proposed by Inoue and Solon (2006) and Born and Breitung (2016).
- For dynamic models or static models with possibly endogenous regressors, the Arellano and Bond (1991) test is widely applied.
  - `estat abond` after `xtabond`, `xtdpdsys`, and `xtdpd`.
  - `xtabond2` and `abar` (Roodman, 2009), the latter for use after `regress`, `ivregress`, `ivreg2`, `newey`, and `newey2`.
  - `estat serial` after `xtdpdbc` (Kripfganz and Breitung, 2022), `xtdpdgmm` (Kripfganz, 2019), and `xtdpdqml` (Kripfganz, 2016).

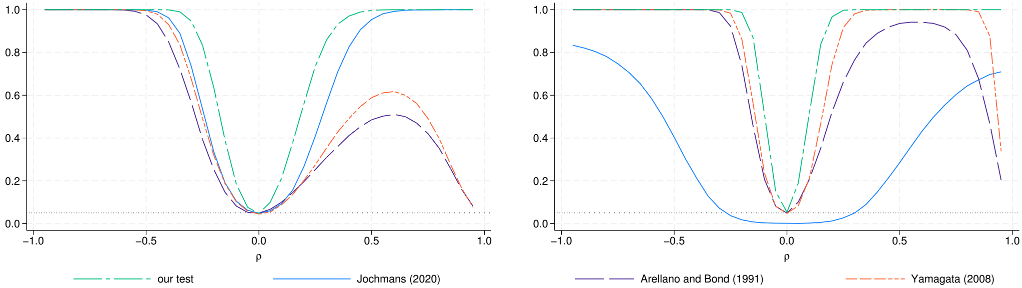
New Stata command xtdpdserial

- Jochmans (2020) proposed a portmanteau test with robustness to heteroskedasticity and better power properties than existing tests (when the time dimension T is very small).
  - xtserialpm (Jochmans and Verardi, 2020): limited to static fixed-effects estimation.
  - In theory, the test is also applicable after (dynamic) models with predetermined or endogenous regressors.
  - The Arellano and Bond (1991) test for no serial correlation at a specified order and the Yamagata (2008) test for no serial correlation at any order, both in first differences, can be regarded as special cases.
- The new command xtdpdserial implements this portmanteau test and a variety of existing and newly proposed special cases (Kripfganz, Demetrescu, and Hosseinkouchack, 2024) for use after regress, xtreg, xtdpdbc, and xtdpdgmm, or as a standalone test for a specified variable.

Power comparison of new and existing tests

Kripfganz, Demetrescu, and Hosseinkouchack (2024)

- Power of serial correlation tests against AR(1) alternatives ($N = 100$):



(a) $T = 5$ (b) $T = 12$

New serial correlation tests

Kripfganz, Demetrescu, and Hosseinkouchack (2024)

- The Jochmans (2020) portmanteau test suffers from a curse of dimensionality when T becomes (moderately) large (similar to the too-many-instruments problem of dynamic panel data GMM estimators).
  - The null hypothesis is $Cov(e_{i,t-s}, \Delta e_{it}) = 0$ for $s \geq 2$ and $s = -1$ (for all t).
- We propose "collapsing" (averaging over time) and "curtailing" (restricting the maximum order of correlation to be tested) for dimensionality reduction.
- Our preferred test with the best power properties is based on "S-differencing" (seasonal or sandwich differencing).
  - The null hypothesis is $Cov(e_{i,t+1} - e_{i,t-s}, \Delta e_{it}) = 0$ for $s \geq 2$ (for all t).
  - Combined with collapsing and/or curtailing, the test retains power when T increases.
  - The test also has power against random-walk alternatives, unlike tests based entirely on first differences.
  - The test does not lose power when $Var(\alpha_i) \to \infty$, unlike the portmanteau test.

New Stata command `xtdpdserial`

- Syntax 1: **xtdpdserial** [*varname*] [*if*] [*in*] [, *options*]

options:		
	pm	portmanteau test: $Cov(\texttt{L\#}.e, \texttt{D}.e) = 0$
	difference	test based on first differences: $Cov(\texttt{L\#D}.e, \texttt{D}.e) = 0$
	sdifference	test based on seasonal differences: $Cov(\texttt{FS(\#+1)}.e, \texttt{D}.e) = 0$
	order(#)	maximum order of serial correlation (in levels)
	lagrange(#$_1$ [#$_2$])	range of lags # to be used
	[full]collapse	collapsed version of the test
	noforward	no forward-looking covariance restrictions
	nobackward	no backward-looking covariance restrictions
	noresiduals	skip check for regression residuals

- Examples:

Jochmans (2020):	`xtdpdserial, pm`
Arellano and Bond (1991):	`xtdpdserial, difference collapse lagrange(2 2)`
Yamagata (2008):	`xtdpdserial, difference collapse`
our test:	`xtdpdserial, sdifference collapse order(2)`

New Stata command xtdpdserial

- Syntax 2: **xtdpdserial** [*varname*] [*if*] [*in*] [, **statistics**(*stats*)]

  stats can be one or more of *stat* [c] [(#$_1$ [#$_2$])]

stat:	**pm**	portmanteau test
	c|fullc	collapsed portmanteau test
	d	test based on first differences
	sd	test based on seasonal differences
optional:	**c**	collapsed version of the test
	(#)	maximum order of serial correlation
	(#$_1$ #$_2$)	range of lags to be used

- Example (same 4 tests as before in a single command line):
  ```
  xtdpdserial, statistics(pm dc(2 2) dc sdc(2))
  ```

- The default is Syntax 2 with **statistics(pm sdc dc fullc)**

Example

```
. webuse abdata                                  // Arellano and Bond (1991) data set

. quietly xtreg n w k, fe vce(robust)            // static FE estimation

. xtdpdserial, statistics(pm dc(2 2) dc sdc(2))

portmanteau test                                      chi2(35)     =    82.6457
H0: no autocorrelation of any order                   Prob > chi2 =     0.0000

collapsed test in first differences                   chi2(1)      =     1.2134
H0: no autocorrelation up to order 3                  Prob > chi2 =     0.2707

collapsed test in first differences                   chi2(6)      =     8.7452
H0: no autocorrelation of any order                   Prob > chi2 =     0.1884

collapsed test in seasonal differences                chi2(1)      =    47.0566
H0: no autocorrelation up to order 2                  Prob > chi2 =     0.0000
```

Example

```
. quietly xtdpdbc n w k, fe lags(1) vce(robust)   // dynamic bias-corrected FE estimation

. xtdpdserial, statistics(pm dc(2 2) dc sdc(2))

portmanteau test                                      chi2(27)     =   54.8003
H0: no autocorrelation of any order                   Prob > chi2 =    0.0012

collapsed test in first differences                   chi2(1)      =    1.2925
H0: no autocorrelation up to order 3                  Prob > chi2 =    0.2556

collapsed test in first differences                   chi2(5)      =   21.7213
H0: no autocorrelation of any order                   Prob > chi2 =    0.0006

collapsed test in seasonal differences                chi2(1)      =    8.9481
H0: no autocorrelation up to order 2                  Prob > chi2 =    0.0028
```

Example

```
. quietly xtdpdbc n w k, fe lags(2) vce(robust)   // dynamic bias-corrected FE estimation

. xtdpdserial, statistics(pm dc(2 2) dc sdc(2))

portmanteau test                                       chi2(20)     =   31.7654
H0: no autocorrelation of any order                    Prob > chi2 =    0.0459

collapsed test in first differences                    chi2(1)      =    0.1563
H0: no autocorrelation up to order 3                   Prob > chi2 =    0.6925

collapsed test in first differences                    chi2(4)      =    3.2948
H0: no autocorrelation of any order                    Prob > chi2 =    0.5098

collapsed test in seasonal differences                 chi2(1)      =    1.3042
H0: no autocorrelation up to order 2                   Prob > chi2 =    0.2534
```

Conclusion

- Serial correlation tests are part of the standard toolkit in applied work.
- Widely used tests require strong assumptions or have low power under certain circumstances.
- Kripfganz, Demetrescu, and Hosseinkouchack (2024) propose new tests with better power properties.
- The new Stata package `xtdpdserial` implements a variety of serial correlation tests robust to heteroskedasticity and valid under deviations from strict exogeneity.
- Word of caution: If the coefficient estimator is inconsistent under the alternative (as typically the case in dynamic models), these tests can be substantially underpowered. Recommended remedy: Use an estimator that is consistent both under the null and alternative hypothesis!

```
ssc install xtdpdserial
net install xtdpdserial, from(http://www.kripfganz.de/stata/)

help xtdpdserial
```