A nonparametric test of separability in structural equations in Stata: `testnp`

Céline Detilleux Jann Goedecke

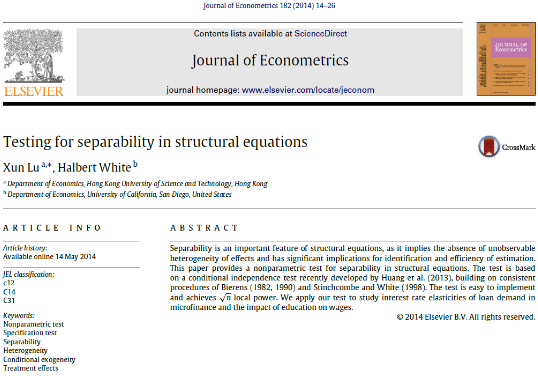
KU Leuven

Belgian Stata Conference, June 2021

Introduction

The paper:

1. review the nonparametric test for separability of structural equations by Lu and White (2014)
2. implement the test in Stata: `testnp`.



Journal of Econometrics 182 (2014) 14–26

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

ELSEVIER

Testing for separability in structural equations

Xun Lu [a,*], Halbert White [b]

[a] Department of Economics, Hong Kong University of Science and Technology, Hong Kong
[b] Department of Economics, University of California, San Diego, United States

ARTICLE INFO

Article history:
Available online 14 May 2014

JEL classification:
c12
C14
C31

Keywords:
Nonparametric test
Specification test
Separability
Heterogeneity
Conditional exogeneity
Treatment effects

ABSTRACT

Separability is an important feature of structural equations, as it implies the absence of unobservable heterogeneity of effects and has significant implications for identification and efficiency of estimation. This paper provides a nonparametric test for separability in structural equations. The test is based on a conditional independence test recently developed by Huang et al. (2013), building on consistent procedures of Bierens (1982, 1990) and Stinchcombe and White (1998). The test is easy to implement and achieves $\sqrt{n}$ local power. We apply our test to study interest rate elasticities of loan demand in microfinance and the impact of education on wages.

Outline

1. The LW separability test
    - The setup
    - Motivations
    - The test statistics
2. The `testnp` command
    - Syntax
    - Options
    - Stored results
3. Simulation exercise
4. Conclusion

The LW separability test

The setup

1. Consider the data generating process:

$$Y = r(X, U) \tag{1}$$

   where
   - Y is the response function of interest
   - X is an observable treatment or cause of interest
   - U denotes the causes that may be unobservable
   - $r()$ is an unknown measurable function

2. Consider a *conditioning instrument* Z, which is potentially correlated with both X and U, but not related to Y:

$$X \perp U|Z \tag{2}$$

The LW separability test

The setup

The LW test of separability is:

$$\boxed{\begin{array}{ll} \mathbb{H}_0: & r(X,U) = r_1(X) + r_2(U) \\ \mathbb{H}_1: & \mathbb{H}_0 \text{ is false} \end{array}}$$

where r_1 and r_2 are unknown measurable functions

- Utility functions with two or more arguments

$$U(c, l) = u(c) + u(l)$$

- Production functions

$$F(K, L^h, L^l) = f(K) + f(L^h, L^l)$$

- In a experiment

$$Y = \alpha + \beta X + \epsilon$$

BUT, separability is typically assumed rather than tested
$\Rightarrow$ empirical test can give valuable insight whether or not it can be justified

- Separability implies that the ME or TE are the same for individual with different unobservable characteristics:

$Y = r(X, U)$	$Y = r_1(X) + r_2(U)$
$\tilde{m}(U) = r(1, U) - r(0, U)$	$\tilde{m}(U) = r_1(1) - r_1(0)$
$m(x, U) = \frac{\partial r(x,U)}{\partial x}$	$m(x, U) = \frac{\partial r_1(x)}{\partial x}$

$$\boxed{\begin{array}{ll} \mathbb{H}_0: & r(X,U) = r_1(X) + r_2(U) \\ \mathbb{H}_1: & \mathbb{H}_0 \text{ is false} \end{array}}$$

BUT, we cannot directly test $\mathbb{H}_0$ since U is unobserved

- LW introduce the random variable $V := Y - E(Y|X,Z)$ and the modified hypothesis can thus be written as:

$$\boxed{\begin{array}{ll} \mathbb{H}_0': & V \perp X|Z \\ \mathbb{H}_1': & \text{Not } \mathbb{H}_0' \end{array}}$$

  where the distribution of V can be entirely expressed in terms of the observable X, Y and Z

- Su and White (2008) show that $\mathbb{H}_0{}'$ is equivalent to:

$$f_Z(z)f_{VXZ}(v,x,z) - f_{XZ}(x,z)f_{VZ}(v,z) = 0 \tag{3}$$

BUT, **slow convergence rate of** $n^{-1/2}h^{-d/4}$

- Following Huang et al. (2016), LW show $\mathbb{H}_0{}'$ is true $\iff \tau(\lambda) = 0$ for any $\lambda \in \Lambda$ where:

$$\tau(\lambda) \equiv \int_{\mathcal{Z}}\int_{\mathcal{V}}\int_{\mathcal{X}} \phi(v,x,z,\lambda)f_Z(z)\cdot[f_{VXZ}(v,x,z) - f_{XZ}(x,z)\cdot f_{VZ}(v,z)]\, dx\, dv\, dz$$

which is consistent and has a **fast convergence rate of** $\sqrt{n}^{-1}$
where
  $\phi(x;\lambda)$ is a *generically comprehensively revealing* function
  λ represents a random vector of dimension $3 + d_z$

- We can re-write $\tau(\lambda)$ in terms of moment conditions:

$$\tau(\lambda) \equiv E_{XVZ}[\phi(X, V, Z; \lambda) f_Z(z)] - E_{VZ}[E_X[\phi(X; V, Z, \lambda)|Z] f_Z(z)] = 0$$

- Using the kernel density estimation, the sample analog of $\tau(\lambda)$ can be written as:

$$\boxed{\begin{aligned}\hat{\tau}_n(\lambda) \equiv \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \Bigg\{ \frac{1}{h_Z^{k_z}} K_z \left(\frac{Z_i - Z_j}{h_Z} \right) \\ \cdot \left[\phi \left(\lambda_0 + \hat{V}_i' \lambda_1 + X_i' \lambda_2 + Z_i' \lambda_3 \right) \right. \\ \left. - \phi \left(\lambda_0 + \hat{V}_i' \lambda_1 + X_j' \lambda_2 + Z_i' \lambda_3 \right) \right] \Bigg\}\end{aligned}}$$

The LW separability test

The test statistics

- The GCR function $\phi(.)$ is the standard normal density function
- λ is a uniform distribution based on the sample X_i, V_i, Z_i with a support interval of length 1
- K_z is a multivariate kernel density estimators:

$$\hat{f}(Z_i) = K_z\left(\frac{Z_i - z}{h_Z}\right) = \prod_{m=1}^{d_z} K\left(\frac{Z_{i,m} - z_m}{h_Z}\right)$$

- K is an univariate kernel density estimators
- V is estimated nonparametically using the Nadaraya-Watson estimator:

$$\hat{V}_i = Y_i - \frac{\sum_{j=1}^{n} K\left(\frac{X_j - X_i}{h_X}\right) K_z\left(\frac{Z_j - Z_i}{h_Z}\right) Y_j}{\sum_{j=1}^{n} K\left(\frac{X_j - X_i}{h_X}\right) K_z\left(\frac{Z_j - Z_i}{h_Z}\right)}$$

- h_X and h_Z are the corresponding bandwidths

- Under the null, $\tau(\lambda) = 0$ should hold for essentially any value of λ (Huang et al., 2016)

- The test statistic is thus expressed as the mean integrated squared $\tau(\lambda)$:

$$\boxed{T_n = \int \hat{\tau}_n(\lambda)^2 f(\lambda) \, d\lambda}$$

- BUT,for finite n, the distribution of $\tau(\lambda)^2$ is nonstandard and difficult to compute
$\Rightarrow$ use a subsampling approach to derive a data driven distribution against which the test statistic can be compared

- Implement the test in Stata version 15 using Mata
- Test can be called by using the command `testnp`

Syntax:

`testnp` *depvar indepvar* [*if*] [*in*], `instruments`*(varlist)* [`bwtype`(*string*) `kerntype`(*string*) `kernorder`(*#*) `intetype`(*string*) `numnodes`(*#*) `accuracy`(*#*) `simtype`(*string*) `numsim`(*#*) `varteststat`(*newvarname*)]

Options:

[`bwtype(`*string*`)` `kerntype(`*string*`)` `kernorder(`*#*`)` `intetype(`*string*`)` `numnodes(`*#*`)` `accuracy(`*#*`)` `simtype(`*string*`)` `numsim(`*#*`)` `varteststat(`*newvarname*`)`]

`kerntype(`*kernel*`)`	the kernel function
`kernorder(`*integer*`)`	the order of the kernel
`bwtype(`*type*`)`	the automatic bandwidth selector
`intetype(`*type*`)`	the integration calculation method
`numnodes(`*#*`)`	the number of nodes
`accuracy(`*#*`)`	the accuracy level
`simtype(`*simulation*`)`	subsampling or bootstrapping
`varteststat(`*newvarname*`)`	the simulated distribution of the test statistic

Scalars	
`r(test_stat)`	Test statistic T_n
`r(p_value)`	p-value
`r(rejection_rate)`	Rejection frequency ($1-p$-value)
`r(kernorder)`	Order of the kernel
`r(numnodes)`	Number of nodes
`r(accuracy)`	Accuracy
`r(numsim)`	Number of simulations used
`r(n)`	Sample size
Locals	
`r(depvar)`	Dependent variable
`r(indepvar)`	Independent (endogenous) variable
`r(instruments)`	Conditioning instruments
`r(kerntype)`	Type of kernel used
`r(bwtype)`	Bandwidth selector used
`r(intetype)`	Method to calculate the integration
`r(simtype)`	Method to empirically derive the asymptotic null distribution
Matrices	
`r(simulation)`	Matrix of simulated distribution of the test statistics

Simulation exercise

We first consider a data-generating process as defined by the following system of equations (DGP1):

$$
\begin{aligned}
\mathbb{H}_0 \text{ is true} &: \quad Y = X + U \\
\mathbb{H}_1 \text{ is true} &: \quad Y = X + U + U \cdot (X + X^2)
\end{aligned}
$$

where

$$
\begin{aligned}
Z &= \varepsilon_Z \\
U &= 0.5 \cdot Z + \sqrt{Z + 4} + 0.5 \cdot \varepsilon_U \\
X &= 0.5 \cdot Z + 2 \cdot \varepsilon_X \\
(\varepsilon_Z, \varepsilon_U, \varepsilon_X) &\sim \mathcal{N}(0, 1)
\end{aligned}
$$

where the sample size is set to n=250

Simulation exercise

```
. // Run the LW nonparametric test for separability
. testnp Y1 X, instruments(Z1 Z2) kern(epa2) bw(silverman) kernor(4) inte(mc) numnodes(100) accuracy(4) sim(boot) numsim(50)
// Stata output
 Kernel type: epa2
 Kernel order: 4
 Integration method: mc
 Number of nodes: 100
 Accuracy: 4
 Simulation type: boot
 Number of simulations: 50
Integrated squared Tau: .006313306
Integrated squared Tau (simulation 1): .000504836, Simulation sample size: 250
Integrated squared Tau (simulation 2): .000185302, Simulation sample size: 250
Integrated squared Tau (simulation 3): .001084876, Simulation sample size: 250
Integrated squared Tau (simulation 4): .007533851, Simulation sample size: 250
Integrated squared Tau (simulation 5): .002812156, Simulation sample size: 250
```

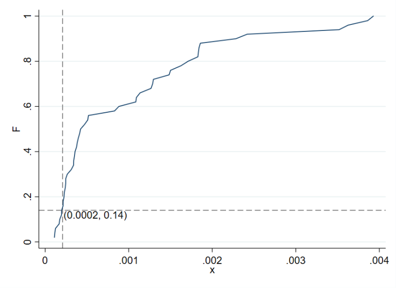
... (*output omitted*) ...

```
Integrated squared Tau (simulation 48): .000735123, Simulation sample size: 250
Integrated squared Tau (simulation 49): .001744335, Simulation sample size: 250
Integrated squared Tau (simulation 50): .00202525, Simulation sample size: 250
H0 rejection rate: .92
```

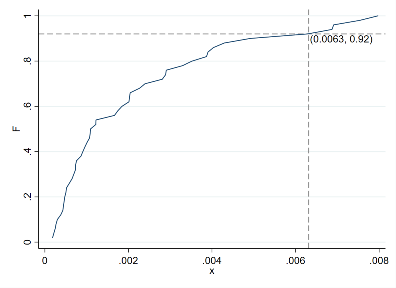
Simulation exercise

Figure: Empirical CDF of test statistic

(a) under $\mathbb{H}_0$



(b) under $\mathbb{H}_1$



Simulation exercise

Table: DGP 1 (n=250, `kerntype`=epa2, `intetype`=mc)

Type of bandwidth selector	silverman	normalscale	oversmoothed	lw
Panel A: p-values (DGP1 under the null)				
Subsample $b=[n^{0.9}]$	0.66	0.88	0.52	0.36
Bootstrap	0.86	0.92	0.64	0.56
Panel B: p-values (DGP1 under the alternative)				
Subsample $b=[n^{0.9}]$	0.00	0.00	0.00	0.00
Bootstrap	0.08	0.02	0.02	0.00

Conclusion

1. This paper reviews the nonparametric test for separability of variables in a structural equation developed by Lu and White (2014)
   - The test's novel feature is that one of the two variables tested for separability is unobserved
   - The test's strength is its feature of consistency and its fast rate of convergence $\sqrt{n}^{-1}$
2. This paper implements the test in Stata by creating the Stata command `testnp`
   - The test statistic is computed entirely in Stata's Mata language
   - Results obtained by two simulation exercises indicate that test delivers good results

References I

Meng Huang, Yixiao Sun, and Halbert White. A flexible nonparametric test for conditional independence. *Econometric Theory*, 32(6): 1434–1482, 2016.

Xun Lu and Halbert White. Testing for separability in structural equations. *Journal of Econometrics*, 182(1):14–26, 2014.

Liangjun Su and Halbert White. A nonparametric hellinger metric test for conditional independence. *Econometric Theory*, 24(04):829–864, 2008.