Diagnostic checking using Subspace Methods

Alfredo García-Hiernaux

Department of Quantitative Economics
Universidad Complutense de Madrid

15th International Conference on Computing in Economics and Finance, Sydney, 2009

Outline

Literature

Since Box and Pierce (1970) many studies have been focused on the portmanteau tests and its properties.

Lately the literature can be split into 2 categories:

1. Papers which relax assumptions of the original test
   - Lobato (2001); Francq et al. (2005)
   - Duchesne and Roy (2004)
2. Works that present improvements in its finite properties
   - Horowitz et al. (2006)

Literature

Since Box and Pierce (1970) many studies have been focused on the portmanteau tests and its properties.

Lately the literature can be split into 2 categories:

1. Papers which relax assumptions of the original test
   - Lobato (2001); Francq et al. (2005)
   - Duchesne and Roy (2004)
2. Works that present improvements in its finite properties
   - Horowitz et al. (2006)

Literature

Since Box and Pierce (1970) many studies have been focused on the portmanteau tests and its properties.

Lately the literature can be split into 2 categories:

1. Papers which relax assumptions of the original test
   - Lobato (2001); Francq et al. (2005)
   - Duchesne and Roy (2004)
2. Works that present improvements in its finite properties
   - Horowitz et al. (2006)

What makes the proposals different from other tests?

- They are obtained by tackling the question from a subspace methods perspective.
- Its asymptotic null distribution is known (no need of computationally expensive simulations)
- They generalize the Box-Pierce's (1970) statistic for single series and the Hosking's (1980) for multivariate processes.
- They are more robust in the presence of outliers that occur in the beginning or the end of the sample.

Brief Introduction to the Subspace Methods

Consider a linear-fixed, stable, strictly minimum-phase system (**Assumption 1**) in innovations form:

$$\left.\begin{aligned} x_{t+1} &= \Phi x_t + E\psi_t \\ z_t &= Hx_t + \psi_t \end{aligned}\right\} Z_f = O_i \overbrace{MZ_p}^{\text{approx } X_f} + V_i \Psi_f$$

$$Z_p = \begin{pmatrix} z_1 & z_2 & \cdots & z_{T-2i+1} \\ z_2 & z_3 & \cdots & z_{T-2i+2} \\ \vdots & \vdots & & \vdots \\ z_i & z_{i+1} & \cdots & z_{T-i} \end{pmatrix}; \quad Z_f = \begin{pmatrix} z_{i+1} & z_{i+2} & \cdots & z_{T-i+1} \\ z_{i+2} & z_{i+3} & \cdots & z_{T-i+2} \\ \vdots & \vdots & & \vdots \\ z_{2i} & z_{2i+1} & \cdots & z_T \end{pmatrix}$$

Ψ_f is as Z_f but with ψ_t instead of z_t. Further $T^* = T - 2i + 1$.

Brief Introduction to the Subspace Methods

Consider a linear-fixed, stable, strictly minimum-phase system (**Assumption 1**) in innovations form:

$$\left.\begin{aligned} x_{t+1} &= \Phi x_t + E\psi_t \\ z_t &= Hx_t + \psi_t \end{aligned}\right\} Z_f = O_i \overbrace{MZ_p}^{\text{approx } X_f} + V_i \Psi_f$$

$$Z_p = \begin{pmatrix} z_1 & z_2 & \cdots & z_{T-2i+1} \\ z_2 & z_3 & \cdots & z_{T-2i+2} \\ \vdots & \vdots & & \vdots \\ z_i & z_{i+1} & \cdots & z_{T-i} \end{pmatrix}; \quad Z_f = \begin{pmatrix} z_{i+1} & z_{i+2} & \cdots & z_{T-i+1} \\ z_{i+2} & z_{i+3} & \cdots & z_{T-i+2} \\ \vdots & \vdots & & \vdots \\ z_{2i} & z_{2i+1} & \cdots & z_T \end{pmatrix}$$

Ψ_f is as Z_f but with ψ_t instead of z_t. Further $T^* = T - 2i + 1$.

Brief Introduction to the Subspace Methods

Consider a linear-fixed, stable, strictly minimum-phase system (**Assumption 1**) in innovations form:

$$\left.\begin{array}{lcl} \boldsymbol{x}_{t+1} & = & \boldsymbol{\Phi x}_t + \boldsymbol{E}\psi_t \\ \boldsymbol{z}_t & = & \boldsymbol{H x}_t + \psi_t \end{array}\right\} \boldsymbol{Z}_f = \boldsymbol{O}_i \overbrace{\boldsymbol{M Z}_p}^{\text{approx } \boldsymbol{X}_f} + \boldsymbol{V}_i \boldsymbol{\Psi}_f$$

$$\boldsymbol{Z}_p = \begin{pmatrix} \boldsymbol{z}_1 & \boldsymbol{z}_2 & \cdots & \boldsymbol{z}_{T-2i+1} \\ \boldsymbol{z}_2 & \boldsymbol{z}_3 & \cdots & \boldsymbol{z}_{T-2i+2} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{z}_i & \boldsymbol{z}_{i+1} & \cdots & \boldsymbol{z}_{T-i} \end{pmatrix}; \quad \boldsymbol{Z}_f = \begin{pmatrix} \boldsymbol{z}_{i+1} & \boldsymbol{z}_{i+2} & \cdots & \boldsymbol{z}_{T-i+1} \\ \boldsymbol{z}_{i+2} & \boldsymbol{z}_{i+3} & \cdots & \boldsymbol{z}_{T-i+2} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{z}_{2i} & \boldsymbol{z}_{2i+1} & \cdots & \boldsymbol{z}_T \end{pmatrix}$$

$\boldsymbol{\Psi}_f$ is as $\boldsymbol{Z}_f$ but with ψ_t instead of $\boldsymbol{z}_t$. Further $T^* = T - 2i + 1$.

Brief Introduction to the Subspace Methods

Consider a linear-fixed, stable, strictly minimum-phase system (**Assumption 1**) in innovations form:

$$\left.\begin{array}{rcl} \boldsymbol{x}_{t+1} & = & \boldsymbol{\Phi x}_t + \boldsymbol{E}\psi_t \\ \boldsymbol{z}_t & = & \boldsymbol{Hx}_t + \psi_t \end{array}\right\} \boldsymbol{Z}_f = \boldsymbol{O}_i \overbrace{\boldsymbol{MZ}_p}^{\text{approx } \boldsymbol{X}_f} + \boldsymbol{V}_i \boldsymbol{\Psi}_f$$

$$\boldsymbol{O}_i = \begin{pmatrix} \boldsymbol{H}' & (\boldsymbol{H\Phi})' & (\boldsymbol{H\Phi}^2)' & \ldots & (\boldsymbol{H\Phi}^{i-1})' \end{pmatrix}'$$

$$\boldsymbol{V}_i = \begin{pmatrix} \boldsymbol{I}_m & 0 & 0 & \ldots & 0 \\ \boldsymbol{HE} & \boldsymbol{I}_m & 0 & \ldots & 0 \\ \boldsymbol{H\Phi E} & \boldsymbol{HE} & \boldsymbol{I}_m & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{H\Phi}^{i-2}\boldsymbol{E} & \boldsymbol{H\Phi}^{i-3}\boldsymbol{E} & \boldsymbol{H\Phi}^{i-4}\boldsymbol{E} & \ldots & \boldsymbol{I}_m \end{pmatrix}$$

Brief Introduction to the Subspace Methods

Consider a linear-fixed, stable, strictly minimum-phase system (**Assumption 1**) in innovations form:

$$\left.\begin{aligned} x_{t+1} &= \Phi x_t + E\psi_t \\ z_t &= H x_t + \psi_t \end{aligned}\right\} Z_f = O_i \overbrace{M Z_p}^{\text{approx } X_f} + V_i \Psi_f$$

$$O_i = \begin{pmatrix} H' & (H\Phi)' & (H\Phi^2)' & \dots & (H\Phi^{i-1})' \end{pmatrix}'$$

$$V_i = \begin{pmatrix} I_m & 0 & 0 & \dots & 0 \\ HE & I_m & 0 & \dots & 0 \\ H\Phi E & HE & I_m & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ H\Phi^{i-2}E & H\Phi^{i-3}E & H\Phi^{i-4}E & \dots & I_m \end{pmatrix}$$

Brief Introduction to the Subspace Methods

- Subspace methods solve the reduced-rank weighted least squares (LS) problem:

$$\boldsymbol{Z_f} = \boldsymbol{O_i M Z_p} + \boldsymbol{V_i \Psi_f} \tag{1}$$

efficiently computed by applying the SVD to $\boldsymbol{W_1 Z_f W_2}$, being

$$\boldsymbol{W_1} = (\boldsymbol{Z_f Z_f'})^{-\frac{1}{2}} \tag{2}$$

$$\boldsymbol{W_2} = \boldsymbol{Z_p'}(\boldsymbol{Z_p Z_p'})^{-1}\boldsymbol{Z_p} \tag{3}$$

First considerations

Assumption 2

Let ψ_t be a sequence of *iid* random variables, with $\mathrm{E}(\psi_t) = 0$ and $\mathrm{E}(\psi_t'\psi_t) = \boldsymbol{Q}$, being $\boldsymbol{Q}$ a positive definite matrix.

Hypothesis

Let H_0 be that there are no correlations up to lag order k and let H_1 be that there exist correlations up to lag order k.

Further:

- For simplicity, we define $i = \mathrm{int}\{(k+1)/2\}$ [1]
- $\boldsymbol{Z}_f = \boldsymbol{OMZ}_p + \boldsymbol{V\Psi}_f \quad \Rightarrow \quad \boldsymbol{Z}_f = \beta \boldsymbol{Z}_p + \boldsymbol{V\Psi}_f$
- β can be estimated as: $\hat{\beta} = \boldsymbol{Z}_f\boldsymbol{Z}_p'(\boldsymbol{Z}_p\boldsymbol{Z}_p')^{-1}$

First considerations

Assumption 2

Let ψ_t be a sequence of *iid* random variables, with $\text{E}(\psi_t) = 0$ and $\text{E}(\psi'_t \psi_t) = \boldsymbol{Q}$, being $\boldsymbol{Q}$ a positive definite matrix.

Hypothesis

Let H_0 be that there are no correlations up to lag order k and let H_1 be that there exist correlations up to lag order k.

Further:

- For simplicity, we define $i = \text{int}\{(k+1)/2\}$ [1]
- $\boldsymbol{Z}_f = \boldsymbol{OMZ}_p + \boldsymbol{V\Psi}_f \quad \Rightarrow \quad \boldsymbol{Z}_f = \boldsymbol{\beta Z}_p + \boldsymbol{V\Psi}_f$
- β can be estimated as: $\hat{\beta} = \boldsymbol{Z}_f \boldsymbol{Z}'_p (\boldsymbol{Z}_p \boldsymbol{Z}'_p)^{-1}$

Exploiting an estimate of β (I)

Proposition 1

Under H_0 and given Assumptions 1 and 2, the covariance matrix of $vec(\hat{\beta}|\mathbf{Z_p})$ can be formulated as $\mathbf{H}^{-1}\mathbf{A}(\mathbf{\Omega} \otimes \mathbf{Q})\mathbf{A}'\mathbf{H}^{-1}$, where $\mathbf{A} = \mathbf{Z}_p \otimes \mathbf{I}_{im}$, $\mathbf{H} = \mathbf{A}'\mathbf{A}$.

In the univariate case, $(m = 1)$ and the noise covariance, Q, is a scalar, Kolmogorov's strong law of large numbers and H_0 ensure that $T_*\mathbf{H}^{-1} \xrightarrow{a.s.} Q^{-1}\mathbf{I}_{i^2}$ and $T_*^{-1}Q\mathbf{A}'\mathbf{\Omega}\mathbf{A} \xrightarrow{a.s.} Q^2\bar{\mathbf{\Pi}}$ and then:

$$\sqrt{T_*}vec(\hat{\beta}|\mathbf{Z}_p) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{\Pi}}).$$

In the multivariate case $(m > 1)$, the same distribution holds if some standardization of the data is carried out.

Exploiting an estimate of β (II)

$$\sqrt{T_*}\, vec(\hat{\beta}|\mathbf{Z}_p) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{\Pi}}).$$

$$\bar{\mathbf{\Pi}} = \begin{pmatrix} \mathbf{I}_{im^2} & \mathbf{\Pi}_{i-1} & \mathbf{\Pi}_{i-2} & \dots & \mathbf{\Pi}_1 \\ \mathbf{\Pi}'_{i-1} & \mathbf{I}_{im^2} & \mathbf{\Pi}_{i-1} & \dots & \mathbf{\Pi}_2 \\ \mathbf{\Pi}'_{i-2} & \mathbf{\Pi}'_{i-1} & \mathbf{I}_{im^2} & \dots & \mathbf{\Pi}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Pi}'_1 & \mathbf{\Pi}'_2 & \mathbf{\Pi}'_3 & \dots & \mathbf{I}_{im^2} \end{pmatrix}_{(im)^2} \qquad (4)$$

with $\mathbf{\Pi}_{i-j} = \begin{pmatrix} \pi_{i-j} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \pi_{i-j} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \pi_{i-j} \end{pmatrix}_{im^2}$ and $\pi_{i-j} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{m(i-j)} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}_{im}$,

where $j = 1, 2, ..., i-1$.

Exploiting an estimate of $\boldsymbol{O}$

A second intuitive idea is to use the information held in an estimate of matrix $\boldsymbol{O}$, defined in (1) ...

... and estimated as $\boldsymbol{W}\hat{\boldsymbol{O}} = \boldsymbol{W}\boldsymbol{Z}_f\boldsymbol{Z}_p'(\boldsymbol{Z}_p\boldsymbol{Z}_p')^{-\frac{1}{2}}$, being $\boldsymbol{W}$ an weighting matrix.

Proposition 2

Under H_0 and assumptions 1 and 2, if $\boldsymbol{W} = (\boldsymbol{Z}_f\boldsymbol{Z}_f')^{-\frac{1}{2}}$, then $\sqrt{T_*}vec(\boldsymbol{W}\hat{\boldsymbol{O}}|\boldsymbol{Z}_p) \xrightarrow{d} N(\boldsymbol{0}, \bar{\boldsymbol{\Pi}})$.

Exploiting an estimate of $\boldsymbol{O}$

A second intuitive idea is to use the information held in an estimate of matrix $\boldsymbol{O}$, defined in (1) ...

... and estimated as $\boldsymbol{W}\hat{\boldsymbol{O}} = \boldsymbol{W}\boldsymbol{Z}_f\boldsymbol{Z}_p'(\boldsymbol{Z}_p\boldsymbol{Z}_p')^{-\frac{1}{2}}$, being $\boldsymbol{W}$ an weighting matrix.

Proposition 2

Under H_0 and assumptions 1 and 2, if $\boldsymbol{W} = (\boldsymbol{Z}_f\boldsymbol{Z}_f')^{-\frac{1}{2}}$, then $\sqrt{T_*}vec(\boldsymbol{W}\hat{\boldsymbol{O}}|\boldsymbol{Z}_p) \xrightarrow{d} N(\boldsymbol{0}, \bar{\boldsymbol{\Pi}})$.

Statistics and its distributions

From the structure of $\bar{\mathbf{\Pi}}$, Prop 3 can be stated:

Proposition 3

For any random matrix $\boldsymbol{A}$ such that $\sqrt{T_*}vec\boldsymbol{A} \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{\Pi}})$, there is an idempotent matrix $\boldsymbol{P}_{(im)^2}$ of rank m^2k, such that:

$$s(\boldsymbol{A}) = T_* vec(\boldsymbol{A})'\boldsymbol{P}vec(\boldsymbol{A}) \xrightarrow{d} \chi^2_{m^2k}.$$

Corollaries

- Under H_0 and being $\hat{\beta} = \bar{\mathbf{Z}}_f\bar{\mathbf{Z}}_p'(\bar{\mathbf{Z}}_p\bar{\mathbf{Z}}_p')^{-1}$, applying Prop 1 gives $s(\hat{\beta}) \xrightarrow{d} \chi^2_{m^2k}$.
- By applying Prop 2, under H_0, then $s(\boldsymbol{W}\hat{\boldsymbol{O}}) \xrightarrow{d} \chi^2_{m^2k}$ holds.

Statistics and its distributions

From the structure of $\bar{\mathbf{\Pi}}$, Prop 3 can be stated:

Proposition 3

For any random matrix $\boldsymbol{A}$ such that $\sqrt{T_*}vec\boldsymbol{A} \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{\Pi}})$, there is an idempotent matrix $\boldsymbol{P}_{(im)^2}$ of rank m^2k, such that:

$$s(\boldsymbol{A}) = T_* vec(\boldsymbol{A})' \boldsymbol{P} vec(\boldsymbol{A}) \xrightarrow{d} \chi^2_{m^2k}.$$

Corollaries

- Under H_0 and being $\hat{\beta} = \bar{\boldsymbol{Z}}_f \bar{\boldsymbol{Z}}'_p (\bar{\boldsymbol{Z}}_p \bar{\boldsymbol{Z}}'_p)^{-1}$, applying Prop 1 gives $s(\hat{\beta}) \xrightarrow{d} \chi^2_{m^2k}$.
- By applying Prop 2, under H_0, then $s(\boldsymbol{W}\hat{\boldsymbol{O}}) \xrightarrow{d} \chi^2_{m^2k}$ holds.

Univariate and multivariate ARMA residuals

One of the most common use of the portmanteau tests is to check the residuals obtained from fitting (V)ARMA models.

But the asymptotic distribution of $s(\hat{\beta})$ and $s(\boldsymbol{W}\hat{\boldsymbol{O}})$ is not as in previous corollaries since the residuals present some linear constraints inherit from the VARMA estimation.

Proposition 4

If $\boldsymbol{z}_t$ (in the SS model) are the residuals from a fitted m-vector ARMA(p, q) model, then, under H_0, $s(\hat{\beta})$ and $s(\boldsymbol{W}\hat{\boldsymbol{O}})$ converge in distribution to a $\chi^2_{m^2(k-p-q)}$.

Univariate and multivariate ARMA residuals

One of the most common use of the portmanteau tests is to check the residuals obtained from fitting (V)ARMA models.

But the asymptotic distribution of $s(\hat{\beta})$ and $s(\boldsymbol{W}\hat{\boldsymbol{O}})$ is not as in previous corollaries since the residuals present some linear constraints inherit from the VARMA estimation.

Proposition 4

If $\boldsymbol{z}_t$ (in the SS model) are the residuals from a fitted m-vector ARMA(p, q) model, then, under H_0, $s(\hat{\beta})$ and $s(\boldsymbol{W}\hat{\boldsymbol{O}})$ converge in distribution to a $\chi^2_{m^2(k-p-q)}$.

Monte Carlo Simulations: Univariate (I)

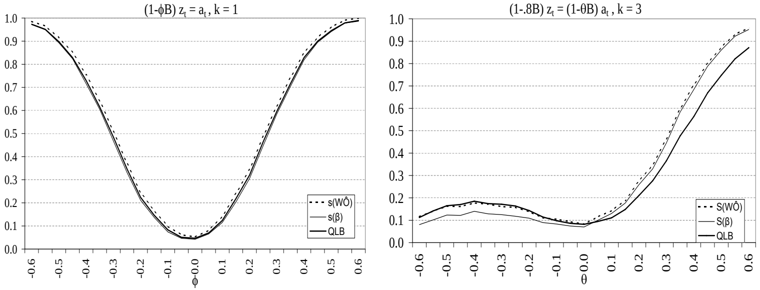

Figure: Empirical power of $s(\boldsymbol{W}\hat{\boldsymbol{O}})$, $s(\hat{\beta})$ and Q_{LB} for a sample size $T = 50$, different ARMA processes and lags (k). The empirical powers are computed with a χ^2_k distribution except for the ARMA(1,1) process where a χ^2_{k-1} is employed as H_0 is tested over the residuals of a misspecified AR(1) model.

Monte Carlo Simulations: Univariate (II)

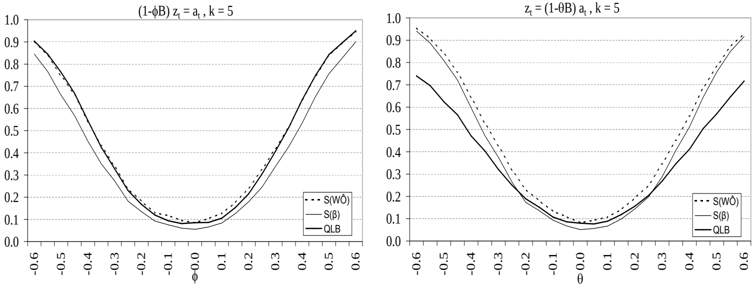

Figure: Empirical power of $s(\boldsymbol{W}\hat{\boldsymbol{O}})$, $s(\hat{\beta})$ and Q_{LB} for a sample size $T = 50$, different ARMA processes and $k = 5$. The empirical powers are computed with a χ^2_k distribution.

Monte Carlo Simulations: Bivariate

$k(\text{lag}) = 1$, $T = 50$, replications $= 5000$.

Table: Empirical size of $s(\boldsymbol{W}\hat{\boldsymbol{O}})$, $s(\hat{\beta})$ and Q_{LB}.

AR	MA	$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	$s(\hat{\beta})$	$\boldsymbol{Q}_{LB}$
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	.049	.046	$\begin{pmatrix} .052 & .055 \\ .057 & .053 \end{pmatrix}$

Table: Empirical power of $s(\boldsymbol{W}\hat{\boldsymbol{O}})$, $s(\hat{\beta})$ and Q_{LB}

AR	MA	$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	$s(\hat{\beta})$	$\boldsymbol{Q}_{LB}$
$\begin{pmatrix} 1 & -.2B \\ 0 & 1-.4B \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	.957	.950	$\begin{pmatrix} .106 & .607 \\ .135 & .732 \end{pmatrix}$
$\begin{pmatrix} 1-.4B & 0 \\ 0 & 1-.8B \end{pmatrix}$	$\begin{pmatrix} 1-.7B & .2B \\ -.3B & 1-.7B \end{pmatrix}$	.196	.187	$\begin{pmatrix} .082 & .093 \\ .105 & .055 \end{pmatrix}$

Monte Carlo Simulations: Bivariate

$k(\text{lag}) = 1$, $T = 50$, replications $= 5000$.

Table: Empirical size of $s(\boldsymbol{W\hat{O}})$, $s(\hat{\beta})$ and Q_{LB}.

AR	MA	$s(\boldsymbol{W\hat{O}})$	$s(\hat{\beta})$	$\boldsymbol{Q}_{LB}$
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	.049	.046	$\begin{pmatrix} .052 & .055 \\ .057 & .053 \end{pmatrix}$

Table: Empirical power of $s(\boldsymbol{W\hat{O}})$, $s(\hat{\beta})$ and Q_{LB}

AR	MA	$s(\boldsymbol{W\hat{O}})$	$s(\hat{\beta})$	$\boldsymbol{Q}_{LB}$
$\begin{pmatrix} 1 & -.2B \\ 0 & 1-.4B \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	.957	.950	$\begin{pmatrix} .106 & .607 \\ .135 & .732 \end{pmatrix}$
$\begin{pmatrix} 1-.4B & 0 \\ 0 & 1-.8B \end{pmatrix}$	$\begin{pmatrix} 1-.7B & .2B \\ -.3B & 1-.7B \end{pmatrix}$	.196	.187	$\begin{pmatrix} .082 & .093 \\ .105 & .055 \end{pmatrix}$

An example with real data (I)

Data

$\boldsymbol{y}_t$: Indices of monthly flour prices in Buffalo, Minneapolis and Kansas City, over the period 08/1972-11/1980.

Transformations: $\boldsymbol{z}_t = \nabla \log \boldsymbol{y}_t$

Table: P-values. H_0 : No correlations up to lag k in $\boldsymbol{z}_t$.

k (lag)	$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	$s(\hat{\beta})$	$\boldsymbol{Q}_{LB}$
1	$.000^*$	$.000^*$	$\begin{pmatrix} .168 & .025^* & .045^* \\ .099 & .026^* & .053 \\ .043^* & .017^* & .063 \end{pmatrix}$
5	$.281$	$.042^*$	$\begin{pmatrix} .813 & .404 & .491 \\ .704 & .408 & .475 \\ .453 & .295 & .533 \end{pmatrix}$

An example with real data (I)

Data

$\boldsymbol{y}_t$: Indices of monthly flour prices in Buffalo, Minneapolis and Kansas City, over the period 08/1972-11/1980.

Transformations: $\boldsymbol{z}_t = \nabla \log \boldsymbol{y}_t$

Table: P-values. H_0 : No correlations up to lag k in $\boldsymbol{z}_t$.

k (lag)	$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	$s(\hat{\beta})$	$\boldsymbol{Q}_{LB}$
1	$.000^*$	$.000^*$	$\begin{pmatrix} .168 & .025^* & .045^* \\ .099 & .026^* & .053 \\ .043^* & .017^* & .063 \end{pmatrix}$
5	$.281$	$.042^*$	$\begin{pmatrix} .813 & .404 & .491 \\ .704 & .408 & .475 \\ .453 & .295 & .533 \end{pmatrix}$

An example with real data (II)

Following $\boldsymbol{Q}_{LB}$ $(k = 1)$ at 5%, a restricted $(\boldsymbol{I} - \boldsymbol{\Phi}_1 B)\boldsymbol{z}_t = \boldsymbol{a}_t$ is estimated:

$$\hat{\boldsymbol{\Phi}}_1 = \begin{pmatrix} 0 & -.188^* & -.035 \\ 0 & -.289^* & 0 \\ -.401^* & .117 & 0 \end{pmatrix} \quad (5)$$

Table: P-values. H_0 : No correlations up to lag k in residuals

Statistic	k (lags) 2	5	10	15
$Q^{\dagger}_{LB}$	.413	.855	.698	.779
$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	.003*	.219	.150	.359
$s(\hat{\beta})$	.000*	.028*	.002*	.000*

$Q^{\dagger}_{LB}$ corresponds to the lowest p-value among all the elements of $\boldsymbol{Q}_{LB}$

An example with real data (II)

Following $\boldsymbol{Q}_{LB}$ $(k = 1)$ at 5%, a restricted $(\boldsymbol{I} - \boldsymbol{\Phi}_1 B)\boldsymbol{z}_t = \boldsymbol{a}_t$ is estimated:

$$\hat{\boldsymbol{\Phi}}_1 = \begin{pmatrix} 0 & -.188^* & -.035 \\ 0 & -.289^* & 0 \\ -.401^* & .117 & 0 \end{pmatrix} \quad (5)$$

Table: P-values. H_0 : No correlations up to lag k in residuals

	k (lags)			
Statistic	2	5	10	15
$Q^{\dagger}_{LB}$	.413	.855	.698	.779
$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	.003*	.219	.150	.359
$s(\hat{\beta})$	.000*	.028*	.002*	.000*

$Q^{\dagger}_{LB}$ corresponds to the lowest p-value among all the elements of $\boldsymbol{Q}_{LB}$

An example with real data (III)

However, if we had followed either $s(\boldsymbol{W}\hat{\boldsymbol{O}})$ or $s(\hat{\beta})$, we would have estimated an unrestricted $(\boldsymbol{I} - \boldsymbol{\Phi}_1 B)\boldsymbol{z}_t = \boldsymbol{a}_t$:

$$\hat{\boldsymbol{\Phi}}_1 = \begin{pmatrix} 1.226^* & -1.355^* & .005 \\ .830^* & -1.027^* & .035 \\ .463 & -.813^* & .142 \end{pmatrix} \quad (6)$$

Table: P-values. H_0: No correlations up to lag k in residuals

	k (lags)			
Statistic	2	5	10	15
$Q^{\dagger}_{LB}$	.447	.729	.642	.744
$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	.944	.953	.611	.704
$s(\hat{\beta})$	.942	.934	.331	.288

$Q^{\dagger}_{LB}$ corresponds to the lowest p-value among all the elements of $\boldsymbol{Q}_{LB}$

An example with real data (III)

However, if we had followed either $s(\boldsymbol{W}\hat{\boldsymbol{O}})$ or $s(\hat{\beta})$, we would have estimated an unrestricted $(\boldsymbol{I} - \boldsymbol{\Phi}_1 B)\boldsymbol{z}_t = \boldsymbol{a}_t$:

$$\hat{\boldsymbol{\Phi}}_1 = \begin{pmatrix} 1.226^* & -1.355^* & .005 \\ .830^* & -1.027^* & .035 \\ .463 & -.813^* & .142 \end{pmatrix} \quad (6)$$

Table: P-values. H_0: No correlations up to lag k in residuals

	k (lags)			
Statistic	2	5	10	15
$Q_{LB}^{\dagger}$	.447	.729	.642	.744
$s(\boldsymbol{W}\hat{\boldsymbol{O}})$	.944	.953	.611	.704
$s(\hat{\beta})$	.942	.934	.331	.288

$Q_{LB}^{\dagger}$ corresponds to the lowest p-value among all the elements of $\boldsymbol{Q}_{LB}$

Summary

- This work tackles the problem of diagnostic checking from the subspace methods point of view.
- Two subspace-based statistics are presented and its asymptotic distributions are derived under H_0.
- Both procedures generalize the Box-Pierce statistic (single series) and the Hoskings' statistic (multivariate case).
- Simulations and an example with real data show that they perform better than the common Q-statistic in different situations.

What's next?

- Trying to extend the tests to models including exogenous variables, e.g., VARMAX or TF.

And finally ...

... THANK YOU FOR LISTENING!

For Further Reading I

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residuals autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332):1509–1526.

Duchesne, P. and Roy, R. (2004). On consistent testing for serial correlation of unknown form in vector time series models. *Journal of Multivariate Analysis*, 89:148–180.

Francq, C., Roy, R., and Zakoïan, J.-M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association*, 100(13):532–544.

Horowitz, J. L., Lobato, I., Nankervis, J. C., and Savin, N. (2006). Bootstrapping the box-pierce q test: A robust test of uncorrelatedness. *Journal of Econometrics*, 133(2):841–862.

For Further Reading II

Lobato, I. (2001). Testing that a dependent process is uncorrelated. *Journal of the American Statistical Association*, 96:1066–1076.