Testing instantaneous causality in presence of non constant unconditional covariance:

Supplementary material

QUENTIN GIAI GIANETTO* AND HAMDI RAÏSSI†

April 29, 2014

1 Technical proofs

Let $\theta_0 = \text{vec}\{(A_{01}, \dots, A_{0p})\}$ and denote by $\hat{\theta}$ a $\sqrt{T}$ -asymptotically normal estimator of θ_0 . Estimators fulfilling such convergence rate in our framework are proposed in Patilea and Raïssi (2012). The residuals are given by $\hat{u}_t = X_t - (\tilde{X}'_{t-1} \otimes I_d)\hat{\theta}$, with $\tilde{X}_{t-1} = (X'_{t-1}, \dots, X'_{t-p})'$.

Proof of (3.1) We first show that the residuals $\hat{u}_t$ and the errors u_t are asymptotically equivalent for our task. Let us define $u_t(\theta) = (u'_{1t}(\theta), u'_{2t}(\theta))' = X_t - (\tilde{X}_{t-1} \otimes I_d)\theta$

*IRMAR-INSA, 20 avenue des buttes de Coësmes, CS 70839, F-35708 Rennes Cedex 7, France.

E-mail: quentin2g@yahoo.fr

†IRMAR-INSA, 20 avenue des buttes de Coësmes, CS 70839, F-35708 Rennes Cedex 7, France.

E-mail: hamdi.raïssi@insa-rennes.fr

for any $\theta \in \mathbb{R}^{pd^2}$. From the Mean Value Theorem we have

$$\begin{aligned} T^{-1} \sum_{t=1}^T \hat{u}_{2t} \otimes \hat{u}_{1t} &= T^{-1} \sum_{t=1}^T u_{2t} \otimes u_{1t} \\ &+ T^{-1} \sum_{t=1}^T \left\{ \frac{\partial u_{2t}(\theta)}{\partial \theta'} \otimes u_{1t}(\theta) + u_{2t}(\theta) \otimes \frac{\partial u_{1t}(\theta)}{\partial \theta'} \right\}_{\theta=\theta^*} (\hat{\theta} - \theta_0), \end{aligned} \quad (1.1)$$

where θ^* is between $\hat{\theta}$ and θ_0 , and $\partial u_t(\theta)/\partial \theta' = -\tilde{X}'_{t-1} \otimes I_d$ is uncorrelated with u_t .

Thus following standard arguments (see Phillips and Xu (2006), Lemma 1), we have

$$T^{-1} \sum_{t=1}^T \hat{u}_{2t} \otimes \hat{u}_{1t} = T^{-1} \sum_{t=1}^T u_{2t} \otimes u_{1t} + o_p(T^{-\frac{1}{2}}).$$

Next we show that

$$\delta_s \Rightarrow \int_0^s (G_2(r) \otimes G_1(r)) dB_{\mathcal{M}}(r) := K(s). \quad (1.2)$$

Recall that under the null hypothesis of no instantaneous causality we have $H_{1t}H'_{2t} = \Sigma_t^{12} = 0$, so that:

$$\begin{aligned} \text{vec}(u_{1t}u'_{2t}) &= \text{vec}(u_{1t}u'_{2t}) - \text{vec}(H_{1t}H'_{2t}) \\ &= H_{2t}\epsilon_t \otimes H_{1t}\epsilon_t - (H_{2t} \otimes H_{1t})\text{vec}(I_d) \\ &= (H_{2t} \otimes H_{1t})\{\epsilon_t \otimes \epsilon_t - \text{vec}(I_d)\}. \end{aligned}$$

Define $v_t^\epsilon := \epsilon_t \otimes \epsilon_t - \text{vec}(I_d)$. We have $E(v_t^\epsilon) = 0$ and $\text{Var}(v_t^\epsilon) = E(\epsilon_t \epsilon'_t \otimes \epsilon_t \epsilon'_t) - \text{vec}(I_d)\text{vec}(I_d)' =: \mathcal{M}$. Therefore from Theorem 3.1 of Hansen (1992) it follows that (1.2) holds. We obtain the result from the Continuous Mapping Theorem, since the functional $f(Y) = \sup_{s \in [0,1]} \|Y(s)\|_2^2$ is continuous for any $Y \in D[0,1]$, the space of càdlàg processes on $[0,1]$. $\square$

Proof of Proposition 1 For the sake of simplicity and with no loss of generality (see (1.1)) let us assume that $X_t = u_t$, so that the error process is observed and there is no autoregressive parameters to estimate. We first show that $\delta_s^{(i)} \Rightarrow^P K(s)$.

Conditionally on the u_t 's, $\delta_s^{(i)}$ is a Gaussian process with mean zero, independent increments and variance

$$E^*(\delta_s^{(i)}\delta_s^{(i)'}) = T^{-1} \sum_{t=1}^{[Ts]} (u_{2t} \otimes u_{1t})(u_{2t} \otimes u_{1t})',$$

where $E^*(.)$ is the expectation under the bootstrap probability measure. The result follows if

$$T^{-1} \sum_{t=1}^{[Ts]} (u_{2t} \otimes u_{1t})(u_{2t} \otimes u_{1t})' \rightarrow \int_0^s (G_2(r) \otimes G_1(r))\mathcal{M}(G_2(r) \otimes G_1(r))'dr,$$

uniformly for all $s \in [0, 1]$. Since $T^{-1} \sum_{t=1}^{[Ts]} (u_{2t} \otimes u_{1t})(u_{2t} \otimes u_{1t})'$ is monotonically increasing and the limit function is continuous, it suffices to establish the pointwise convergence following (Hansen (2000), proof of Lemma A.10). This holds using similar arguments to those used in Patilea and Raïssi (2012), Lemmas 7.3 and 7.4. From the Continuous Mapping Theorem

$$\sup_{s \in [0,1]} \|\delta_s^{(i)}\|_2^2 \Rightarrow^P \sup_{s \in [0,1]} \|K(s)\|_2^2,$$

and hence we obtain the desired result. $\square$

2 The effect of lag length selection prior testing for instantaneous causality

The residuals of the model are correlated if the autoregressive order of the VAR process is chosen too small. In that case, all the tests considered in the paper may be subject to misleading outputs. On the other hand if a too large lag length is selected, a large number of unnecessary parameters are introduced in the model which can lead to a loss of efficiency.

In view to choose a relevant lag length, the smaller autoregressive order delivering uncorrelated residuals is generally selected. Classical Box-Pierce or Ljung-Box portmanteau tests are often used to find this order. The Box-Pierce test is proposed in Chitturi (1974) and the Ljung-Box test is developed in Hosking (1980) for VAR models with i.i.d. errors. It is important to underline that these tests are not valid in our framework of time-varying unconditional covariance. In that case, the portmanteau tests developed in Patilea and Raïssi (2013) can be used. Indeed such goodness-of-fit tests ensure that the probability of selecting a too small lag length tends to zero, while the probability of selecting a lag length greater than the true order is adequately controlled as the sample size is increased.

By way of illustration, let us consider the following bivariate VAR(2) model :

$$\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.2 \\ 0.3 & -0.3 \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0.1 & 0.3 \\ 0.1 & 0.4 \end{pmatrix} \begin{pmatrix} X_{1,t-2} \\ X_{2,t-2} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \quad (2.1)$$

where the innovations are Gaussian with variance structure $\Sigma(r)$ fulfilling the assumption **A1**. Two cases are considered for this structure :

- **Case 1: Setting under the null.** It does not exist an instantaneous causality relation between $X_{1,t}$ and $X_{2,t}$:

$$\Sigma(r) = \begin{pmatrix} \Sigma^{11}(r) & 0 \\ 0 & \Sigma^{22}(r) \end{pmatrix} \quad \forall r \in (0, 1].$$

where $\Sigma^{11}(r) = \Sigma^{22}(r) = 1 + 11 \times r$ correspond to the non constant variances of the innovations' components.

- **Case 2: Setting under the alternative.** It exists an instantaneous causality relation between $X_{1,t}$ and $X_{2,t}$:

$$\Sigma(r) = \begin{pmatrix} \Sigma^{11}(r) & \Sigma^{12}(r) \\ \Sigma^{12}(r) & \Sigma^{22}(r) \end{pmatrix} \quad \forall r \in (0, 1].$$

where $\Sigma^{12}(r) = \Sigma^{21}(r) = 3 \times \sin(2\pi r)$ is such that $\int_0^1 \Sigma^{12}(r)dr = 0$ with $\Sigma^{12}(r) \neq 0$ almost everywhere on $(0, 1]$, and $\Sigma^{11}(\cdot)$, $\Sigma^{22}(\cdot)$ are defined as in Case 1.

For a given sample size, 1000 time series are generated. For each time series, the minimum autoregressive order which is not rejected by the considered portmanteau test is selected. In our experiments, a significant level of 5% is used in the portmanteau tests and we stop the lag length selection at $p_{max} = 6$. Recall that the true autoregressive order is $p = 2$.

Next the instantaneous causality tests W_{st} , W_w and W_b are performed with this selected autoregressive order. In Table 1 and 2, the Patilea and Raïssi test is used for the prior lag length selection. The relative rejection frequencies under the null of no instantaneous causality are displayed in Table 1 and the relative rejection frequencies under the alternative are displayed in Table 2. Table 3 and Table 4 display these results when the classical Ljung-Box portmanteau test is used to select the lag length. However let us recall that this test is not valid in presence of time-varying covariance. In order to evaluate the impact of the choice of an autoregressive order on the instantaneous causality tests, Table 4 and Table 5 display respectively the rejection frequencies under the null and under the alternative when the true autoregressive order $p = 2$ is known.

For the bootstrap W_b test and the White W_w test we observe that the results become similar as the sample size is increased in the three cases (lag length selected by the Patilea-Rassi and the Ljung-Box goodness-of-fit tests, or with a known lag length). We only remark some additional rejections for small samples but the general conclusion remains true: the tests based on the spurious assumption of constant covariance suffer from a severe loss of power while the bootstrap W_b seems to provide satisfactory results. It also emerges that the standard test W_{st} has rejections

frequencies under the null somewhat far from the nominal level whether the true lag length is known or not, which illustrates the inadequacy of the standard test under the null hypothesis of no instantaneous causality as outlined in Section 2 of the paper.

		Asymptotic nominal level								
		1%			5%			10%		
		W_{st}	W_w	W_b	W_{st}	W_w	W_b	W_{st}	W_w	W_b
Sample size	50	0.028	0.021	0.019	0.090	0.084	0.079	0.145	0.143	0.136
	100	0.079	0.063	0.066	0.179	0.162	0.160	0.240	0.219	0.226
	200	0.037	0.026	0.026	0.098	0.080	0.085	0.141	0.127	0.130
	500	0.018	0.012	0.014	0.069	0.045	0.042	0.132	0.112	0.119
	1000	0.020	0.008	0.013	0.083	0.059	0.059	0.134	0.103	0.107

Table 1: The relative rejection frequencies under the null for the studied tests with asymptotic nominal level 1%, 5%, 10% and a chosen lag length with the Patilea-Rassi test.

		Asymptotic nominal level								
		1%			5%			10%		
		W_{st}	W_w	W_b	W_{st}	W_w	W_b	W_{st}	W_w	W_b
Sample size	50	0.037	0.020	0.021	0.075	0.064	0.062	0.121	0.103	0.113
	100	0.083	0.050	0.044	0.200	0.164	0.166	0.270	0.239	0.243
	200	0.045	0.018	0.027	0.165	0.105	0.151	0.237	0.181	0.326
	500	0.036	0.014	0.535	0.091	0.046	0.942	0.149	0.097	0.996
	1000	0.025	0.005	0.995	0.076	0.040	1.000	0.160	0.078	1.000

Table 2: The relative rejection frequencies under the alternative for the studied tests with asymptotic nominal level 1%, 5%, 10% and a chosen lag length with the Patilea-Rassi test.

		Asymptotic nominal level								
		1%			5%			10%		
		W_{st}	W_w	W_b	W_{st}	W_w	W_b	W_{st}	W_w	W_b
Sample size	50	0.023	0.017	0.019	0.096	0.088	0.084	0.145	0.126	0.128
	100	0.039	0.025	0.024	0.120	0.101	0.093	0.166	0.160	0.165
	200	0.019	0.011	0.015	0.064	0.054	0.061	0.120	0.100	0.113
	500	0.012	0.004	0.012	0.081	0.063	0.069	0.139	0.111	0.117
	1000	0.018	0.011	0.011	0.070	0.047	0.041	0.126	0.085	0.087

Table 3: The relative rejection frequencies under the null for the studied tests with asymptotic nominal level 1%, 5%, 10% and a chosen lag length with the Ljung-Box test.

		Asymptotic nominal level								
		1%			5%			10%		
		W_{st}	W_w	W_b	W_{st}	W_w	W_b	W_{st}	W_w	W_b
Sample size	50	0.055	0.033	0.039	0.128	0.104	0.106	0.156	0.154	0.153
	100	0.088	0.047	0.043	0.217	0.170	0.176	0.318	0.293	0.287
	200	0.050	0.021	0.020	0.139	0.095	0.121	0.224	0.175	0.313
	500	0.029	0.007	0.536	0.080	0.046	0.947	0.153	0.095	0.997
	1000	0.027	0.004	0.995	0.083	0.041	1.000	0.140	0.081	1.000

Table 4: The relative rejection frequencies under the alternative for the studied tests with asymptotic nominal level 1%, 5%, 10% and a chosen lag length with the Ljung-Box test.

		Asymptotic nominal level								
		1%			5%			10%		
		W_{st}	W_w	W_b	W_{st}	W_w	W_b	W_{st}	W_w	W_b
Sample size	50	0.011	0.006	0.007	0.042	0.029	0.028	0.061	0.061	0.060
	100	0.019	0.013	0.012	0.065	0.049	0.046	0.114	0.105	0.116
	200	0.015	0.009	0.008	0.081	0.062	0.061	0.131	0.101	0.112
	500	0.020	0.013	0.013	0.060	0.040	0.038	0.135	0.110	0.102
	1000	0.020	0.013	0.009	0.076	0.052	0.053	0.134	0.100	0.104

Table 5: The relative rejection frequencies under the null for the studied tests with asymptotic nominal level 1%, 5%, 10% and a known lag length $p = 2$.

		Asymptotic nominal level								
		1%			5%			10%		
		W_{st}	W_w	W_b	W_{st}	W_w	W_b	W_{st}	W_w	W_b
Sample size	50	0.074	0.046	0.047	0.189	0.162	0.167	0.234	0.221	0.235
	100	0.100	0.055	0.056	0.250	0.209	0.201	0.339	0.308	0.319
	200	0.051	0.022	0.020	0.148	0.086	0.125	0.229	0.170	0.323
	500	0.024	0.008	0.527	0.078	0.044	0.945	0.156	0.097	0.996
	1000	0.026	0.008	0.995	0.084	0.041	1.000	0.146	0.093	1.000

Table 6: The relative rejection frequencies under the alternative for the studied tests with asymptotic nominal level 1%, 5%, 10% and and a known lag length $p = 2$.

References

- Chitturi, R. (1974). Distribution of residual autocorrelations in multiple autoregressive schemes. *Journal of the American Statistical Association*, 69(348):928–934.
- Hansen, B. (1992). Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory*, 8(04):489–500.
- Hansen, B. (2000). Sample splitting and threshold estimation. *Econometrica*, 68(3):575–603.
- Hosking, J. (1980). The multivariate portmanteau statistic. *Journal of the American Statistical Association*, 75(371):602–608.
- Patilea, V. and Raïssi, H. (2012). Adaptive estimation of vector autoregressive models with time-varying variance: application to testing linear causality in mean. *Journal of Statistical Planning and Inference*, 142(11):2891–2912.
- Patilea, V. and Raïssi, H. (2013). Corrected portmanteau tests for var models with time-varying variance. *Journal of Multivariate Analysis*, 116:190–207.
- Phillips, P. and Xu, K.-L. (2006). Inference in autoregression under heteroskedasticity. *Journal of Time Series Analysis*, 27(2):289–308.