

Volume 35, Issue 2

Step-by-Step Causality Revisited: Theory and Evidence

Konstantinos N. Konstantakis
National Technical University of Athens

Panayotis G. Michaelides
National Technical University of Athens

Abstract

The present paper introduces a VAR model with exogenous variables for testing one-sided (non-)causality by extending the works of Dufour and Renault (1998) and Dufour et al. (2006). In this context, it derives a test statistic for formally investigating one sided (non-)causality, while providing a simple algorithm for implementing the one sided (non-)causality test in a system framework and not equation-by-equation extending, thus, Dufour et al. (2006). We illustrate our approach by using a monthly dataset including dummy variables on Total Car Sales in the area of Athens over the period 2003-2012. According to our findings all variables cause the evolution of Total Sales cycles immediately and for almost eight (8) quarters when most of the causality effects die out completely. Clearly, future research on extending the methodology to a panel set-up would be of great interest.

We would like to thank the anonymous referees and the Editors for their diligent reading of the manuscript. The second author (P.G.M.) gratefully acknowledges the support provided by the Siemens-IKY Fellowship of Excellence.

Citation: Konstantinos N. Konstantakis and Panayotis G. Michaelides, (2015) "Step-by-Step Causality Revisited: Theory and Evidence", *Economics Bulletin*, Volume 35, Issue 2, pages 871-877

Contact: Konstantinos N. Konstantakis - konstantinos.konstantakis@gmail.com, Panayotis G. Michaelides - pmichael@central.ntua.gr.

Submitted: March 03, 2015. **Published:** April 09, 2015.

1. Introduction

In a seminal paper in *Econometrica*, Dufour and Renault (1998) introduced the notion of *step-by-step* or *short-run* causality based on the idea that two time series X_t and Y_t could interact in a causal scheme via a third variable Z_t . In this context, despite the fact that X_t could not cause Y_t one period ahead it could cause Z_t one period ahead i.e. Z_{t+1} , and Z_{t+1} could cause Y_t two periods ahead i.e. Y_{t+2} . Therefore, $X_t \rightarrow Y_{t+2}$, even though $X_t \nrightarrow Y_{t+1}$.

In order to investigate the timing pattern of causality, Dufour et al. (2006) in the *Journal of Econometrics*, extended the work of Dufour and Renault (1998) by considering a class of VAR (p) models in different horizons h. Their choice for considering a VAR scheme was based on the bi-direction of causality. Of course, there are cases when we are interested only in one sided (non-)causality, e.g. in order to account for the recent global crisis. In such case, a dummy variable would have to be used to capture the impact of the recent global crisis on other variables of interest, e.g. local ones, such as the Total Car Sales in Greece. However, we have no serious reason to believe that the Total Car Sales in Greece and/or any other local variables of interest could have any causal predictive ability, even in the short run, on the global recession.

In other words, this means that the dummy variable used to capture the recent global recession should not be incorporated in the VAR model proposed by Dufour et al. (2006). It should rather be incorporated in an extended model in the form of an exogenous variable i.e. in simple words it should appear only in the right hand side of the block of VAR equations. Needless to say, this has serious implications for the test statistic that was proposed by Dufour et al. (2006) which is constructed to be bi-directional. Hence, a variable acting as exogenous would render the symmetric test statistic proposed by Dufour et al. (2006) meaningless.

In the meantime, the choice of Dufour et al. (2006) to estimate the VAR model using equation-by-equation OLS instead of SURE or 2S-GLS is inappropriate when the error terms are correlated across different equations, as Dufour et al. (2006, p. 346) themselves point out. In this work, we will set out a methodology for testing how one sided (non-)causality can be tested using a VAR (p) scheme augmented by an exogenous (set of) variable(s) in cases we are interested only in one sided causality between the variables, using 2S-GLS estimator which accounts for the possible error terms correlation across different equations.

In brief, the paper contributes to the literature as follows: (i) It introduces a relevant VAR model with exogenous variables for testing one sided non-causality accounting for the possibility of dummy variables; (b) it derives a test statistic for formally investigating one sided non-causality; (c) it provides a simple algorithm for implementing one sided non-causality using 2S-GLS estimator which accounts for the possible error terms correlation across different equations; and (d) it illustrates this technique using a monthly dataset (2000-2012) on Total Car Sales in the area of Athens, Greece which was hit severely by the recent recession.

2. Methodology

Remark 1: In what follows, we illustrate how one sided (non-)causality can be tested using a VAR (p) scheme augmented by an exogenous set of variables in cases we are interested only in one sided causality between the variables.

2.1 Formulation of one sided non-causality

Here, we set out the one sided causality testing method taking into consideration the case where both dummy and quantitative time series variables are employed.

Consider the following VAR (p) model augmented by exogenous dummy and/or quantitative variables:

$$Y_t = a + \sum_{k=1}^p \pi_k Y_{t-k} + \sum_{q=0}^Q \beta_q D_{t-q} + u_t \quad (1)$$

where: Y_t is an (1xm) vector of variables; a is a (1xm) vector of constant terms; D_t is a vector of (1x1) qualitative (dummy) or quantitative variables and u_t is a (1xm) vector of error terms such that $E(u_t u_s) = \sigma_{ii} I$ if $t = s$ and $E(u_t u_s) = \sigma_{ij} I$ if $t \neq s$, where I is the identity matrix.

Note that the exogenous variables D_t ought to have a lag structure in order to be able to properly apply the concept of short-run causality.

Remark 2: Extending the work by Dufour et al. (2006), we propose an estimation strategy which accounts for the fact that the various disturbances might be contemporaneously correlated, due the same set of regressors that account for the exogenous variables.

Following Dufour et al. (2006), the model described in (1) corresponds to horizon $h=1$. In order to test for the existence of non-causality in horizon h , a model of the following form is considered:

$$Y_{t+h} = a^{(h)} + \pi^{(h)} Y_{t,p} + \beta^{(h)} D_{t,q} + u_{t+h}^{(h)} \quad (2)$$

where: $Y_{t,p} = (Y_t, Y_{t-1}, \dots, Y_{t-p+1})$, $\pi^{(h)} = (\pi_1^{(h)}, \dots, \pi_p^{(h)})$, $\beta^{(h)} = (\beta_0^{(h)}, \beta_1^{(h)}, \dots, \beta_q^{(h)})$ and $u_{t+h}^{(h)} = (u_{1,t+h}^{(h)}, \dots, u_{m,t+h}^{(h)})$ for $t=1, \dots, T-h$ and $h < T$.

Equation (2) can be written in matrix form as:

$$Y_{t+h} = \Gamma X + u \quad (3)$$

where $Y_{t+h} = [Y_{1,t+h}, \dots, Y_{m,t+h}]$ is a (1xm) vector which denotes the m-quantitative variables that enter the model; $X = [I_T; Y_{1,t-1}, \dots, Y_{1,t-p}; \dots; Y_{m,t-1}, \dots, Y_{m,t-p}; D_{1,t-1}, \dots, D_{1,t-q}; \dots; D_{l,t-1}, \dots, D_{l,t-q}]$ is an $(2m+1) \times \max\{t-p+1, t-q+1\}$ matrix that includes both quantitative and qualitative variables; $\Gamma = [a_1, \dots, a_m; \pi_{1,1}, \dots, \pi_{1,p}; \dots; \pi_{m,1}, \dots, \pi_{m,p}; \beta_0, \dots, \beta_{0,q}; \dots; \beta_l, \dots, \beta_{l,q}]$ is the inverse of a $(2m+1) \times [\max\{p, q+1\}]$ matrix of coefficients and $u = [u_{1,t+h}, \dots, u_{m,t+h}]$ is a (1xm) vector of idiosyncratic shocks such that $u \sim N(0, \Sigma)$ so that the variance covariance matrix is of the form: $\Omega = \Sigma \otimes I$, where $\Sigma = (\sigma_{ij})$ and I the identity matrix, with $\det(\Omega) \neq 0$.

Proposition 1: (*Asymptotic normality of GLS in a stationary VAR (p, h)*)

Any VAR (p, h) model described in (2) that can be written in the following form, is asymptotically normally distributed:

$$Y_{t+h} = \Gamma X + u,$$

Where $u \sim N(0, \Omega)$ and the variance covariance matrix is of the form: $\Omega = \Sigma \otimes I$, where $\Sigma = (\sigma_{ij})$ and I the identity matrix, with $\det(\Omega) \neq 0$ and $\frac{1}{T} X'X \rightarrow_{T \rightarrow \infty}^p \Delta_p$ with $\det(\Delta_p) \neq 0$.

Proof: It is a straightforward application of the sketch provided in Dufour et al. (2006, p. 343) (Proposition 1) using GLS estimation instead of LS.

2.2 Distribution of the test statistic for non-causality at horizon h

For a given horizon \bar{h} , we need to test the hypothesis that: $H_0^{(\bar{h})}: D_i \nrightarrow Y_{jt} / I(D_i)$ i.e. the i-th dummy variable does not cause in horizon h the j-th quantitative variable.

Theorem 1: (*Asymptotic distribution of the test criterion for one-sided non-causality at horizon h in a VAR (p) augmented by exogenous quantitative/qualitative variables*)

Under Proposition 1 and the assumption that:

$$H_{0_{D_i \rightarrow Y_{jt}/I(D_i)}}^{(\bar{h})}: R\Gamma^{(\bar{h})} = r \text{ in } Y_{t+h} = \Gamma X + \mathbf{u}$$

then: $V(\widehat{\Gamma^{(\bar{h})}}) \xrightarrow{p}_{T \rightarrow \infty} \Delta_p^{-1} V_{ip} \Delta_p^{-1}$ and is distributed as follows:

$$\mathcal{D}\left(H_0^{(\bar{h})}\right) = T[R\Gamma^{(\bar{h})} - r]' [R'\Delta_p^{-1}\Omega^{-1}\Delta_p R][R\Gamma^{(\bar{h})} - r] \sim \chi^2(\max\{p, q+1\}).$$

Proof: In equation (2) we need to test $H_{0_{D_i \rightarrow Y_{jt}/I(D_i)}}^{(\bar{h})}: \beta_i^{(\bar{h})} = 0$ given that $\forall h \in \{1, \dots, \bar{h}-1\}$ it holds that $\beta_i^{(h)} = 0$, which in turn yields:

$$H_{0_{D_i \rightarrow Y_{jt}/I(D_i)}}^{(\bar{h})}: R\Gamma^{(\bar{h})} = r \quad (4)$$

where: $R = [0, \dots, 0_m; 0, \dots, 0_{2mp}; 0, \dots, 1_i, \dots, 0_{l_x(q+1)}]$

Now, we have that the GLS estimator $\widehat{\Gamma^{(\bar{h})}}$, for $\Gamma^{(\bar{h})}$ is:

$$\widehat{\Gamma^{(\bar{h})}} = \Gamma^{(\bar{h})} + (\mathbf{X}'\Omega^{-1}\mathbf{X})^{-1}\mathbf{X}'\Omega^{-1}\mathbf{u}$$

Hence:

$$\sqrt{T}(\widehat{\Gamma^{(\bar{h})}} - \Gamma^{(\bar{h})}) = \left(\frac{1}{T}\mathbf{X}'\Omega^{-1}\mathbf{X}\right)^{-1} \frac{1}{\sqrt{T}}\mathbf{X}'\Omega^{-1}\mathbf{u}$$

Under standard regularity conditions (White 1999):

$$\sqrt{T}(\widehat{\Gamma^{(\bar{h})}} - \Gamma^{(\bar{h})}) \xrightarrow{L}_{T \rightarrow \infty} N\left(0, V(\widehat{\Gamma^{(\bar{h})}})\right) \quad (5)$$

with: $\det(V(\widehat{\Gamma^{(\bar{h})}})) \neq 0$.

Remark 3: The $V(\widehat{\Gamma^{(\bar{h})}})$ can be consistently estimated using the Newey-West heteroskedasticity and autocorrelation consistent covariance (HAC) matrix estimator extending Dufour et al. (2006, p. 346) who suggested using it without however implementing it:

$$V(\widehat{\Gamma^{(\bar{h})}}) = HAC = \widehat{Q}_0 + \sum_{j=1}^k w(j, k)(\widehat{Q}_j + \widehat{Q}_j')$$

where: $\widehat{Q}_j = \frac{1}{T} \sum_{t=j+1}^T X_t \mathbf{u}_t \mathbf{u}_{t-j}' X_{t-j}'$, $\forall j = 1, \dots, k$

and $w(j, k)$ is a lag window, and k is the lag truncation parameter.

$$\widehat{V}_T(\widehat{\Gamma^{(\bar{h})}}) \xrightarrow{p}_{T \rightarrow \infty} V(\widehat{\Gamma^{(\bar{h})}})$$

Now, suppose that $\frac{1}{T}\mathbf{X}'\mathbf{X} \xrightarrow{p}_{T \rightarrow \infty} \Delta_p$ with $\det(\Delta_p) \neq 0$, and let:

$$V_{ip} = Var\left(\frac{1}{\sqrt{T}}\mathbf{X}'\Omega^{-1}\mathbf{u}/\mathbf{X}\right) = \frac{1}{T} Var(\mathbf{X}'\Omega^{-1}\mathbf{u}/\mathbf{X}) \Leftrightarrow$$

$$V_{ip} = Var\left(\frac{1}{\sqrt{T}}\mathbf{X}'\Omega^{-1}\mathbf{u}/\mathbf{X}\right) = \frac{1}{T}\mathbf{X}'Var(\Omega^{-1}\mathbf{u}/\mathbf{X})\mathbf{X} \quad (6)$$

Therefore, it is easy to infer that:

$$\text{Var}\left[\left(\frac{1}{T}X'\Omega^{-1}X\right)^{-1}\frac{1}{\sqrt{T}}X'u\right] = \Delta_p^{-1}V_{ip}\Delta_p^{-1} \quad (7)$$

Combining equations (7) and (5) we get that:

$$V(\Gamma^{\widehat{h}}) \xrightarrow{p}_{T \rightarrow \infty} \Delta_p^{-1}V_{ip}\Delta_p^{-1} \quad (8)$$

Meanwhile, in order to test for non-causality of the quantitative/qualitative variables that enter as exogenous in the augmented VAR (p) model, at a given horizon h, we propose the following modified algorithm which builds on Dufour *et al.* (2006).

Step1: An augmented VAR model as in equation (3) is fitted for using GLS estimation and the Newey-West heteroskedasticity and autocorrelation consistent covariance (HAC) for horizon $h=1$ and we obtain the estimates $\widehat{\pi}_\kappa$, $\widehat{\beta}_m$ and $\widehat{\Omega}$.

Step 2: A restricted augmented VAR model using GLS estimation as described in equation (4) is fitted and we obtain the estimates $\pi^{(h)}$ and $\beta^{(h)}$.

Step 3: We compute the test statistic \mathcal{D} for testing non-causality at horizon h i.e. we test the hypothesis $H_{0,D_i \rightarrow Y_{jt}/I(D_i)}^{(h)}: \beta_{im} = 0, m = 0, 1, \dots, M, i \in \{1, \dots, l\}, j \in \{1, \dots, m\}$. We denote $\mathcal{D}_0^{(h)}$ the test statistic based on actual data.

Step 4: We draw N simulated samples from equation (4) using Monte Carlo with $\pi^{(h)} = \pi^{(h)}$, $\beta^{(h)} = \beta^{(h)}$ and $\Omega = \widehat{\Omega}$. We impose the constraints of non-causality at horizon h i.e. $\beta_{im} = 0, m = 0, 1, \dots, M, i \in \{1, \dots, l\}, j \in \{1, \dots, m\}$ and we compute the test statistic for non-causality at horizon h, i.e. $\mathcal{D}_n^{(h)}, n \in \{1, \dots, N\}$.

Step 5: We compute the simulated p-values based on the following formula:

$$\hat{p}_N[x] = \{1 + \sum_{n=1}^N I[\mathcal{D}_n^{(h)} - x]\} / (N + 1)$$

Step 6: We reject the null hypothesis of non-causality at horizon h i.e. $H_{0,D_i \rightarrow Y_{jt}/I(D_i)}^{(h)}$, at level α if $\hat{p}_N[\mathcal{D}_0^{(h)}] \leq \alpha$.

In what follows, we apply the proposed methodology for testing short run causality effects of a number of macroeconomic and dummy variables on the cyclical component of Car Sales in the area of Athens, Greece, which was severely hit by the recent recession.

2. Empirical Analysis

The Greek crisis has reached points that are directly comparable only to the Great Recession including an approximate 20% contraction of GDP in the period 2008-2013 and a very high unemployment rate equal to 27%. The car sales sector is an important industry for the Greek economy since it accounts for a significant part of government revenues, especially through the registration taxes that are directly implemented whenever a car sale takes place as well as through the presumptions implemented once a year. The car sales sector in Greece was significantly affected by the ongoing crisis with a reduction of total sales that exceeded 20%, which in turn affected government revenues. Hence, it is of great importance to investigate the step-by-step predictive ability of the various factors on the car sales industry fluctuations over the last 13 years, using monthly data.

3.1 Data and Variables

The data used are monthly for the period 2000-2012. The data regarding Total Car Sales in the Area of Athens come from AMVIR (Association of Motor Vehicle Importers Representatives); Unemployment and GDP come from the Hellenic Statistical Authority (EL.STAT), while the data on Fuel prices come from the Observatory of Fuel Prices. All quantitative variables in the model are in constant 2005 prices in millions €.

In what follows, we make use of the following notation: $TScycle_t$ is the cyclical component of Total car sales in Athens, extracted by means of Baxter King Filtering; $GDPcycle_t$ is the cyclical component of Greek GDP extracted by means of Baxter King Filtering; UN_t is the local unemployment rate; GDP_t is the Greek GDP; F_t is the fuel price; C_t is the dummy variable of the global recession taking the value 1 in the time interval (2006 (M4)-2012 (M12)) and 0 elsewhere; P_t is the dummy variable of presumptions taking the value 1 in the time period 2009 (M5)-2009 M(8) and 0 elsewhere; RT_t is the dummy variable of the registration taxes taking the value of 1 in the period 2004 (M1) - 2008 (M12) and 0 elsewhere and L_t is the dummy variable of the loans directed to the car market taking the value 1 over the period 2003 (M1)- 2008 (M12) and 0 elsewhere.

3.2 Econometric estimation

We start by examining the stationarity characteristics of the time series. According to Table I, the majority of time series variables were found to be non-stationary, except for GDPcycle and Car Total Sales cycle that were expected to be found stationary, as filtered time series. Nevertheless, all variables exhibit stationarity in first differences (Table II). In this context, all variables with the exception of the cyclical variables are regarded to be integrated of degree one i.e. I(1).

Table I: ADF test original variables			Table II: ADF test first differences		
Variable	p-value	Stationarity	Variable	p-value	Stationarity
GDP	0.36	No	GDP	0	Yes
Unemployment	0.99	No	Unemployment	0.04	Yes
Fuel price	0.59	No	Fuel price	0.01	Yes
TScycle	0	Yes			
GDPcycle	0.03	Yes			

In the presence of I(1) variables we have to examine the existence of cointegrating relationships. To this end, Table III presents the results of Johansen's test.

Table III: Johansen Cointegration Test

Maxrank	LogLikelihood	Eigenvalue	Trace Statistic	CriticalValues	Cointegration
0	-2490.57		156.69	47.21	No
1	2461.04	0.34	97.61	29.68	
2	2435.42	0.3	46.39	15.41	
3	2418.09	0.22	11.73	3.76	
4	2412.23	0.08			

The results indicate that there is no cointegration among the variables therefore we proceed with studying the timing pattern of causality. Before proceeding to the non-causality tests we examined the time horizon, i.e. the maximum lag length of the VAR model using AIC (Table IV).

Table IV: Lag length selection using Akaike Information Criterion (AIC)

Lag	LL	df	p-value	AIC
9	-1954.16	16	0.01	32.77
10	-1934.77	16	0.01	32.72
11	-1899.3	16	0.01	32.42
12	-1836.17	16	0	31.69
13	-1826.45	16	0.24	31.88
14	-1815.17	16	0.13	31.93
15	-1794.31	16	0.05	31.97

According to Table IV, twelve (12) lags were selected as the optimum. In this context, we proceed by testing for one sided non-causality for an horizon of twelve (12) periods based on the methodology presented earlier using 10,000 bootstrapped replications. The results are presented in TableV.

Table V: Step-by-step causality results

<i>RT_t does not cause TSCycle_t</i>			<i>PT_t does not cause TSCycle_t</i>			<i>L_t does not cause TSCycle_t</i>		
Lag	χ^2	p-value	Lag	χ^2	p-value	Lag	χ^2	p-value
1	314.41	0	1	315.15	0	1	314.96	0
2	36.13	0	2	36.36	0	2	36.12	0
3	0.88	0.64	3	1.63	0.44	3	1.17	0.55
4	9.48	0	4	11.6	0	4	10.35	0
5	7.43	0.01	5	10.66	0	5	9.39	0
6	6.32	0.02	6	6.56	0.02	6	8.88	0
7	4.26	0.05	7	3.42	0.04	7	7.35	0
8	1.32	0.35	8	4.52	0.05	8	4.44	0.04
9	0.99	0.44	9	1.44	0.33	9	2.15	0.12
10	0.88	0.56	10	1.01	0.42	10	1.51	0.35
11	0.76	0.66	11	0.95	0.48	11	0.79	0.66
12	0.75	0.68	12	0.89	0.52	12	0.69	0.75

<i>C_t does not cause TSCycle_t</i>			<i>UN_t does not cause TSCycle_t</i>			<i>F_t does not cause TSCycle_t</i>		
Lag	χ^2	p-value	Lag	χ^2	p-value	Lag	χ^2	p-value
1	313.95	0	1	316.11	0	1	326.46	0
2	35.93	0	2	36.47	0	2	43.61	0
3	1.58	0.45	3	0.728	0.69	3	33.25	0
4	12.38	0	4	9.52	0	4	19.79	0
5	11.32	0	5	7.45	0.01	5	12.25	0
6	10.75	0	6	12.45	0	6	4.65	0.03
7	4.65	0.03	7	11.62	0	7	4.32	0.05
8	4.44	0.04	8	13.25	0	8	4.29	0.05

9	1.63	0.24	9	2.16	0.11	9	1.79	0.25
10	1.49	0.32	10	1.56	0.22	10	1.66	0.28
11	0.9	0.42	11	1.62	0.17	11	1.59	0.33
12	0.92	0.39	12	1.55	0.21	12	0.82	0.49
<i>GDPcycle_t does not cause TSCycle_t</i>			<i>GDP_t does not cause TSCycle_t</i>					
Lag	χ^2	p-value	Lag	χ^2	p-value			
1	314.28	0	1	455.47	0			
2	35.36	0	2	90.53	0			
3	0.46	0.79	3	47.05	0			
4	11.70	0	4	50.02	0			
5	10.58	0	5	4.66	0.02			
6	4.33	0.05	6	4.44	0.04			
7	1.68	0.23	7	1.49	0.32			
8	0.97	0.32	8	0.78	0.65			
9	0.88	0.44	9	0.69	0.78			
10	0.32	0.85	10	0.53	0.88			
11	0.12	0.91	11	0.10	0.95			
12	0.09	0.96	12	0.06	0.99			

The results of the short run causality tests (Table V) suggest that all variables cause the evolution of Total Sales cycles immediately (i.e. the p-value is approximately equal to 0), and for almost eight (8) quarters when most of the causality effects die out completely (i.e. the p-value is greater than 0.10, at the 10% level).

4. Conclusion

The present paper introduced a VAR model with exogenous variables for testing one sided non-causality by extending the works of Dufour and Renault (1998) and Dufour et al. (2006). In this context, it derived a test statistic for formally investigating one sided non-causality, while providing a simple algorithm for implementing the one sided non-causality test in a system framework and not equation by equation through OLS extending, thus, Dufour et al. (2006). We illustrated our approach by using a monthly dataset including dummy variables on Total Car Sales in the area of Athens over the period 2000-2012. According to our findings all variables cause the evolution of Total Sales cycles immediately and for almost eight (8) quarters when most of the causality effects die out completely. Clearly, future research on extending the methodology to a panel set-up would be of great interest.

References

- Dufour, J.-M., Renault, E. (1998), Short-run and long-run causality in time series: theory, *Econometrica*, 66 (5), 1099–1125.
- Dufour J.-M. and Pelletier D. and Renault E. (2006), Short run and Long run Causality in Time series: Inference, *Journal of Econometrics*, 132 (2), 337-362.
- White, H. (1999), *Asymptotic Theory for Econometricians*, Academic Press, San Diego, Florida.