ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES

ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES **01-2018**

Dealing with Endogeneity in Threshold Models using Copulas: An Illustration to the Foreign Trade Multiplier

By

Dimitris Christopoulos, Peter McAdam and Elias Tzavalis

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5 / Fax: 210 8238249
76, Patission Street, Athens 104 34 Greece.
Tel.: (+30) 210 8203303-5 / Fax: (+30) 210 8238249
E-mail: econ@aueb.gr / www.aueb.gr

Dealing with Endogeneity In Threshold Models using Copulas: An Illustration to the Foreign Trade Multiplier*

Dimitris Christopoulos[a], Peter McAdam[b] and Elias Tzavalis[a]

[a] Athens University of Economics & Business; [b] DG Research, European Central Bank

February 9, 2018

Abstract

We suggest a new method dealing with the problem of endogeneity of the threshold variable in single regression threshold models and seemingly unrelated systems of them based on copula theory. This theory enables us to relax the assumption that the threshold variable is normally distributed and to capture the dependence between the error term and the threshold variable in each regime of the model independently of the marginal distribution of the threshold variable. This distribution can be estimated non-parametrically conditionally on the value of threshold parameter. To estimate the slope and threshold parameters of the model adjusted for the endogeneity of the threshold variable, we suggest a two-step concentrated least squares estimation method where the threshold parameter is estimated based on a search procedure, in the first step. A Monte Carlo study indicates that the suggested method deals with the endogeneity problem of the threshold variable satisfactorily. As an empirical illustration, we estimate a threshold model of the foreign-trade multiplier conditional on the real exchange rate volatility regime. We suggest a bootstrap procedure to examine if there are significant differences in the foreign-trade multiplier effects across the two regimes of the model, under potential endogeneity of the threshold variable.
JEL classification: C12, C13, C21, C22
Keywords: Threshold model, SUR systems, Copulas, Kourtellos et al. (2016), foreign trade multiplier.

*We thank George Kapetanios, Dimitris Karlis, Andrew Patton and Peter Phillips for comments on an earlier version of the paper.

1 Introduction

There has recently been growing interest in the econometric literature on modelling nonlinear relationships based on threshold models. This class of models can capture the effects of different regimes of the economy on structural, or reduced form, econometric relationships based on an observable variable, which plays the role of the threshold variable. They can be also employed to estimate endogenously from the data the value of the threshold parameter splitting the sample into at least two regimes, by treating this parameter as an unknown quantity. On this front, they can shed light on critical levels of observed variables with economic policy interest, above (or below) which regime switching occurs.

Examples of applications of threshold models and relationships in economics (see Teräsvirta et al. (2010) for a survey) include studies on business cycles (e.g., Potter (1995)), the debt-growth relationship (e.g., Reinhart and Rogoff (2010), Kourtellos et al. (2016)), financial markets and volatility (e.g., Gospodinov (2005), Griffin et al. (2007), Tsay (2010) for a survey), monetary policy models (e.g., Davig and Leeper (2007), Kazanas and Tzavalis (2015)). The econometrics of threshold models have been developed in a number of recent studies (e.g., Hansen (1999, 2000), Gonzalo and Pitarakis (2002), Caner and Hansen (2004) and Pitarakis (2008), Kapetanios (2001, 2010) among others).

One problem that has recently attracted interest is that of the endogeneity of the threshold variable, implied by its contemporaneous correlation with the structural error term. Ignoring this will lead to biased estimates of the threshold and slope parameters of the models. To overcome it, Kourtellos et al. (2016), in an influential work, suggested augmenting the structural threshold regression with the inverse Mills ratio terms. These terms can capture the contemporaneous correlation between the error term and the threshold variable, appearing in each regime of the model. The Kourtellos et al. approach gains its methodological insights from the limited dependent variable literature (see Heckman (1979)), assuming endogenous sample selection. An important point to note, though, is that the approach relies on the assumption that both the error term and threshold variable are normally distributed.

In this paper, we relax the normality assumption for the threshold variable. As one would expect, there is likely to exist a plethora of economic variables which can be considered as threshold variables and for which the normality assumption is violated. To control for the effects of the endogeneity of the threshold variable on the parameter estimates of the model, including the threshold parameter, we rely on the copula theory. This theory enables us to specify the joint distribution of the regression error term and the threshold variable based on a copula function capturing the dependence between them and allowing for their marginal distributions to not be necessarily of the same family.[1] This allows for a much greater flexibility in specifying the joint density of the structural error term and threshold variable. When the copula is Gaussian and the marginal distribution of the error term is normal, we can analytically derive a linear relationship between the error term and a copula-type transformation of the threshold variable in each regime of the model. Based on this relationship, we can adjust the threshold model for the endogeneity effects of the threshold variable, by simply adding into the right hand side (rhs) of the model the copula transformation of the threshold variable. The same can be done if both the copula and the error term are members of the elliptical family of distributions.

Alternative methods to deal with the non-normality of the threshold variable were recently proposed by Kourtellos et al. (2017) and Yu and Phillips (2017), who proposed semi-parametric and nonparametric estimators of the threshold parameter, respectively. Our method avoids the challenges of the semi-parametric and nonparametric estimators of the threshold or the remaining parameters of the model and, at the same time, it relies on a simple linear regression method of capturing the endogenous threshold variable bias effects on the estimates of these parameters.

The paper extends the above, dealing with the problem of endogeneity of the threshold variable, to a seemingly unrelated regression (SUR) framework and it suggests consistent estimation and inference procedures for the slope and threshold parameters. This estimation procedure is in accordance with that of Caner and Hansen (2004), and Kourtellos et al. (2016): it relies on a two-step concentrated least squares estimation method of the slope parameters of the model where in the first step the threshold parameter value is consistently estimated based on a search procedure. Conditional on the estimate of the threshold parameter, we can obtain consistent es-

[1]For a concise overview of applications of copula theory to economic and financial data, see, e.g., Patton (2006), Karlis and Nikoloulopoulos (2009), Patton (2012), and Fan and Patton (2014). Park and Gupta (2012) used the copula method to capture endogeneity of a regressor.

timates of the slope parameters of the model, in the second step. The performance of the suggested estimation procedure to satisfactorily estimate the true value of the threshold parameter, which is of major interest in our study, is evaluated by a Monte Carlo (MC) study. There, we consider alternative distributions of the structural error term and the threshold variable, such as the normal and Student-t distributions.[2]

We then implement the method to estimate a threshold model of the foreign trade multiplier, quantifying the effects of exports on real output growth, conditional on the level of the real exchange rate volatility which stands in for the threshold variable. A value of the multiplier above (below) unity means that an increase in exports as a percentage of real output will increase output more (less) than the initial export increase. Ambiguous evidence on the significance of this multiplier effects on growth may be attributed to ignoring regime shifts in real exchange rate and its volatility capturing differences between flexible and stable (or pegging) real effective exchange rate regimes. As recently noted (see, e.g., Tenreyro (2007)), real exchange rate volatility may impart negative, or positive, effects on exports and the real economic growth rate. Almost certainly, though, it can be treated as an endogenous variable in a threshold model estimating the foreign trade multiplier. The negative effects of this variables on growth can be attributed to the cost of real exchange rate volatility discouraging investment (see, e.g., Grier and Smallwood (2013)). Alternatively, any positive effects can be attributed to the fact that, in order to compensate for persistent real exchange volatility, exporters increase their export volumes and supply. In so doing, they maximize their profits and/or revenues.

In our empirical analysis, we use data for the G7 countries (Canada, France, Italy, Japan, UK, US and Netherlands) over the annual period 1963 to 2015.[3] As a measure of real exchange rate, we employ the real effective exchange rate which takes into account the trade partners' effects and it can smooth out favorable and unfavorable exchange rate movements due to bilateral exchange rate volatility. We suggest a bootstrap testing procedure to formally examine if there are significant differences in the foreign trade multiplier effects across the two regimes of the model, under potential endogeneity of the threshold variable. Our results suggest that, for most countries

[2]For the threshold variable, we also considered a gamma distribution. Results available on request.

[3]We substituted The Netherlands for Germany in the 'G7' given reunification-induced breaks in the data. The sample size was chosen to have a common sample across the countries.

examined, the multiplier effects depends on the exchange rate volatility regime. For countries like the US, UK, Japan and France, the foreign trade multiplier exceeds unity when the exchange rate volatility is at its high volatility regime. This result is also supported by SUR based estimates of the threshold model, allowing for heterogeneity in the parameters of the model.

The rest of the paper is organized as follows. **Section 2** presents our method to deal with the problem of endogeneity in a single equation framework and discusses the estimation and inference procedures of the model adjusted by the copula transformations of the threshold variable. **Section 3** extends the model to the SUR framework. **Section 4** carries out the MC exercise. **Section 5** implements the method to estimate the foreign trade multiplier, which constitutes our empirical illustration. **Section 6** concludes.

2 Model setup and the Copula method

Consider the following threshold model:

$$y_{it} = x'_{it}\beta^{(1)}\mathcal{I}(z_{it} \leq \delta) + x'_{it}\beta^{(2)}\mathcal{I}(z_{it} > \delta) + \varepsilon_{it}, \tag{1}$$

where $t = 1, 2, ..., T$ and $i = 1, 2, .., N$ denote the time and units (equations) of the model, $x_{it} = (1, x_{2it}, ...x_{Kit})'$ is a $(K \times 1)$ vector of independent variables, $\mathcal{I}(A)$ is an indicator function with $\mathcal{I}(A) = 1$ if event $A = \{z_{it} \leq \delta\}$ occurs, and zero otherwise, where z_{it} is an observable economic variable which constitutes the threshold variable of the model and δ is the threshold parameter. $\beta^{(h)} = (\beta_1^{(h)}, \beta_2^{(h)}, ...\beta_K^{(h)})'$ is the vector of slope coefficients of the explanatory variables of the model, collected in vector x'_t, at the two distinct regimes considered by the model, denoted as $h = \{1, 2\}$, and ε_{it} is a $NIID\left(0, \sigma^2_{\varepsilon_i}\right)$ error term, for all i. In our analysis, we assume that N is fixed, while T grows large.

Let us assume that threshold variable z_{it} is endogenous, which means that it is contemporaneously correlated with the error term ε_{it}, implying that $\mathbb{E}(z_{it}\varepsilon_{it}) \neq 0$ and $\mathbb{E}(\varepsilon_{it}|z_{it}) \neq 0$, for all i and t. To see how the endogeneity of z_{it} affects the estimates of the slope coefficients of model (1), assume for the moment that x'_t is exogenous, i.e.,

$\mathbb{E}(\varepsilon_{it}|x_{kit}) = 0$, for all regressors k. Then, equation (1) implies

$$\mathbb{E}(y_{it}|x_{it}, z_{it}) = \begin{cases} x'_{it}\beta^{(1)} + \mathbb{E}(\varepsilon_{it}|z_{it} \leq \delta) & \text{if} \quad z_{it} \leq \delta \\ \\ x'_{it}\beta^{(2)} + \mathbb{E}(\varepsilon_{it}|z_{it} > \delta) & \text{if} \quad z_{it} > \delta \end{cases} \tag{2}$$

for all i, since $\mathbb{E}(\varepsilon_{it}|x_{it}, z_{it}) = \begin{cases} \mathbb{E}(\varepsilon_{it}|z_{it} \leq \delta) & \text{if} \quad z_{it} \leq \delta \\ \\ \mathbb{E}(\varepsilon_{it}|z_{it} > \delta) & \text{if} \quad z_{it} > \delta \end{cases}$ The last relationship indicates that, in order to consistently estimate the threshold parameter δ and the vector of the slope parameters the model $\beta^{(h)}$ in a single regression framework (i.e., for $N = 1$), we must control for the effects of conditional expectation terms $\mathbb{E}(\varepsilon_{it}|z_{it} \leq \delta)$ and $\mathbb{E}(\varepsilon_{it}|z_{it} > \delta)$ on the conditional mean $\mathbb{E}(y_{it}|x_{it}, z_{it})$ in each regime. Consistent estimation of the above parameters requires that $\mathbb{E}(y_{it}|x_{it}, z_{it})$ is correctly specified. To this end, Kourtellos et al. (2016) suggested adding to the rhs of model (1) closed forms of terms $\mathbb{E}(\varepsilon_{it}|z_{it} \leq \delta)$ and $\mathbb{E}(\varepsilon_{it}|z_{it} > \delta)$, given by the inverse Mills ratio terms. These are obtained under the assumption that ε_{it} and z_{it} are normally distributed.

Copulas enables us to relax the normality assumption for z_{it} and control for threshold variable endogeneity effects on the parameter estimates of model (1) by adding into its rhs copula-type transformation of z_{it} obtained by the cumulative distribution functions (cdf) of the truncated probability density functions (pdfs) of z_{it}, which are defined in each regime of the model as $f(z_{it}|z_{it} \leq \delta)$ and $f(z_{it}|z_{it} > \delta)$. This can be justified by the converse of Sklar's theorem, implying that we can combine any two univariate distributions with a copula, capturing the dependence between ε_{it} and z_{it}. Gaussian, or Student-t, copulas, which are part of the elliptical family of copulas, enables us to derive a linear relationship between ε_{it} and the copula transformation of z_{it}, under the two different regimes of the threshold model.

More specifically, using copulas the joint probability densities $f(\varepsilon_{it}, z_{it}|z_{it} \leq \delta)$ and $f(\varepsilon_{it}, z_{it}|z_{it} > \delta)$ can be written, respectively, as follows:

$$\begin{aligned} f(\varepsilon_{it}, z_{it}|z^{(1)}_{it}) &\equiv f(\varepsilon_{it}, z_{it}|z_{it} \leq \delta) = c^{(1)}(U_{\varepsilon_{it}}, U_{z^{(1)}_{it}})f(\varepsilon_{it})f(z_{it}|z_{it} \leq \delta) \\ f(\varepsilon_{it}, z_{it}|z^{(2)}_{it}) &\equiv f(\varepsilon_{it}, z_{it}|z_{it} > \delta) = c^{(2)}(U_{\varepsilon_{it}}, U_{z^{(2)}_{it}})f(\varepsilon_{it})f(z_{it}|z_{it} > \delta) \end{aligned} \tag{3}$$

where $z_{it}^{(1)}$ and $z_{it}^{(2)}$ denote, respectively, truncated random variables $z_{it}^{(1)} = z_{it} \times \mathcal{I}(z_{it} \leq \delta)$ and $z_{it}^{(2)} = z_{it} \times \mathcal{I}(z_{it} > \delta)$, $c^{(1)}(U_{\varepsilon_{it}}, U_{z_{it}^{(1)}})$ and $c^{(2)}(U_{\varepsilon_{it}}, U_{z_{it}^{(2)}})$ are copula functions, where $U_{\varepsilon_{it}}$ and $U_{z_{it}^{(h)}}$, are uniformly $(0,1)$ distributed random variables. In particular, $U_{\varepsilon_{it}}$, $U_{z_{it}^{(1)}}$ and $U_{z_{it}^{(2)}}$ constitute probability integral transformations (PITs) of ε_{it}, $z_{it}^{(1)}$ and $z_{it}^{(2)}$, respectively, which use the cdfs of ε_{it}, $z_{it}^{(1)}$ and $z_{it}^{(2)}$. The truncated pdfs $f(z_{it}|z_{it}^{(1)})$ and $f(z_{it}|z_{it}^{(2)})$, employed in (3), also constitute probability distributions, which integrate to one, when appropriately scaled by the value of the cdf of $f(z_t)$ at δ (e.g. Green (2009), appendix B).

Under a bivariate Gaussian copula, $c^{(h)}(U_{\varepsilon_{it}}, U_{z_{it}^{(h)}})$ can be written as

$$c^{(h)}(U_{\varepsilon_{it}}, U_{z_{it}^{(h)}}) = \phi^{(h)}(\varepsilon_{it}^{*}, z_{it}^{*h}) \tag{4}$$

where $\varepsilon_{it}^{*} = \Phi^{-1}(U_{\varepsilon_{it}})$ and $z_{it}^{*(h)} = \Phi^{-1}(U_{z_{it}^{(h)}})$ are known as copula-transformed variables, $\phi^{(h)}(\cdot)$ is the bivariate standard Gaussian pdf and $\Phi(\cdot)$ is the Gaussian cdf and $\rho^{(h)}$ is the correlation coefficient between ε_{it}^{*} and $z_{it}^{*(h)}$. Using (4) and (3), the conditional pdf $f(\varepsilon_{it}, z_{it}|z_{it}^{(h)})$, is given as

$$f(\varepsilon_{it}|z_{it}^{(h)}) = \frac{f(\varepsilon_{it}, z_{it}|z_{it}^{(h)})}{f(z_{it}|z_{it}^{(h)})} = c^{(h)}(U_{\varepsilon_{it}}, U_{z_{it}^{(h)}}) f(\varepsilon_{it}). \tag{5}$$

After some algebra, $f(\varepsilon_{it}|z_{it}^{(h)})$ can be analytically written as follows:

$$f(\varepsilon_{it}|z_{it}^{(h)}) = \frac{1}{\sigma_{\varepsilon_i}\sqrt{1-\rho^{(h)}}\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \frac{\left(\varepsilon_{it} - \sigma_{\varepsilon_i}\rho^{(h)} z_{it}^{*(h)}\right)^2}{\sigma_{\varepsilon_i}^2 \sqrt{1-(\rho^{(h)})^2}} \right\}. \tag{6}$$

Proof: Appendix A

The formula of pdf given by (6) implies that there exist a linear relationship between ε_{it} and the copula-transformed variables $z_{it}^{*(h)}$, for $h = \{1, 2\}$, given by,

$$\varepsilon_{it} = \sigma_{\varepsilon_i}\rho^{(h)} z_{it}^{*(h)} + \sigma_{\varepsilon_i}\sqrt{1-(\rho^{(h)})^2} e_{it}^{(h)},$$

where $e_{it}^{(h)}$ is an $NIID(0, 1)$ error term. The last relationship implies that $\mathbb{E}(\varepsilon_{it}|z_{it} \leq \delta) = \sigma_{\varepsilon_i}\rho^{(1)} z_{it}^{*(1)}$ and $\mathbb{E}(\varepsilon_{it}|z_{it} > \delta) = \sigma_{\varepsilon_i}\rho^{(2)} z_{it}^{*(2)}$, given that the copula-transformed variables $z_{it}^{*(1)}$ and $z_{it}^{*(2)}$constitute transformations of the truncated random variables $z_{it}^{(1)}$ and $z_{it}^{(2)}$, respectively. Given this relationship, we can employ the following reduced form equation of model (1) to estimate its slope parameters:

$$y_{it} = (x'_{it}\beta^{(1)} + \lambda^{(1)} z^{*(1)}_{it})\mathcal{I}(z_{it} \le \delta) + (x'_{it}\beta^{(2)} + \lambda^{(2)} z^{*(2)}_{it})\mathcal{I}(z_{it} > \delta) + e_{it}, \quad (7)$$

where $\lambda^{(h)} = \sigma_{\varepsilon_i}\rho^{(h)}$, e_{it} is a $NIID(0,1)$ error term defined above across the two regimes.[4] Model (7) can be employed to consistently estimate threshold parameter δ and the vector of slope parameters $\beta^{(h)}$, for $h = \{1,2\}$, controlling for possible endogeneity effects of threshold variable z_{it} on δ and $\beta^{(h)}$. The values of the copula-transformed variables $z^{*(h)}_{it}$, entered into the rhs of (7), can be calculated by integrating the truncated pdfs $f(z_{it}|z_{it} \le \delta)$ and $f(z_{it}|z_{it} > \delta)$, for a given value of δ. Then, we can obtain $z^{*(h)}_{it} = \Phi^{-1}(U_{z^{(h)}_{it}})$, based on its copula transformation definition.

Two remarks can be made for the copula based method adjusting the threshold regression model (1) for the endogeneity of its threshold variable, presented above:

1. The method can be easily applied to the case that scalar threshold variable z_{it} is given as $z_{it} = w'_{it}\pi + u_{it}$, where w_{it} is a vector of exogenous variables and $\mathbb{E}(\varepsilon_{it}|u_{it}) \neq 0$ (see Kourtellos et al. (2016)). In this case, expectation terms $\mathbb{E}(\varepsilon_{it}|z_{it} \le \delta)$ and $\mathbb{E}(\varepsilon_{it}|z_{it} > \delta)$ can be respectively written as $\mathbb{E}(\varepsilon_{it}|u_{it} \le \delta - w'_{it}\pi)$ and $\mathbb{E}(\varepsilon_{it}|u_{it} > \delta - w'_{it}\pi)$, which can be replaced by their copula transformations in augmented model (7). As is argued in the next section, the above reduced form of z_{it}, implying that u_{it} can also play the role of the threshold variable of the model, may prove very convenient in applied work when model (1) is estimated in the SUR framework and there are common or correlated covariates of z_{it} across i.

2. The method can be straightforwardly extended to control for possible endogeneity of any explanatory variable of model (1), collected in vector x_t. This extension involves the inclusion of copula transformations of the explanatory variables in the rhs of (7), denoted as x^*_{kit}, for $k = 1, 2 \ldots K$, see Park and Gupta (2012). These variables can be calculated by PITs of regressors x_{kit}. Note that this result holds, directly, by the definition of trivariate copula $c^{(h)}(U_{\varepsilon_{it}}, U_{x_{kit}}, U_{z^{(h)}_{it}})$. Alternatively, given estimates of the threshold parameter δ we can employ instrumental variables, or 2SLS, estimation procedures to deal with the problem

[4]Note that a similar relationship can be derived when ε_{it} is Student-t distributed and a Student-t copula is used.

of endogenous explanatory variables of model (1), in a second step – see, for instance, Canner and Hansen (2004) and Rothfelder and Bodea (2016).

2.1 Estimation Methodology

In this section, we describe in more detail a consistent method of estimating model (7) in a single equation framework (i.e., for $N = 1$). This method is based on the two-step concentrated least squares (LS) method suggested by Caner and Hansen (2004), and Kourtellos et al. (2016) for threshold models.

In the first step, we estimate the threshold parameter δ based on a search procedure. Conditional on the estimate of δ, in the next step, we can estimate the remaining parameters of the model adjusted by the copula-transformed variables $z_{it}^{*(h)}$, collected in vector $\theta(\delta) = (\beta^{(1)}, \beta^{(2)}, \lambda^{(1)}, \lambda^{(2)})'$.

More specifically, δ can be estimated by solving the following non-linear LS optimization problem:

$$\hat{\delta} = \arg\min_{\delta \in Q_z} RSS(\delta),$$

where $RSS(\delta) = \sum_{t=1}^{T} \hat{e}_{it}^2$ is the residual sum of squares of (7) and Q_z is the set of observable (sample) values of z_{it}. As in Kourtellos et al. (2016), due to overlaps, we include in the rhs of the model both copula-transformed variables $z_{it}^{*(1)}$ and $z_{it}^{*(2)}$ in the above estimation procedure independently of the regimes of the model for each δ. The set Q_z is chosen to efficiently estimate the threshold parameter δ, by leaving a sufficient number of observations in each of the two regimes of the model. The values of variables $z_{it}^{*(h)}$, included in the rhs of (7), are generated as noted before, by integrating pdfs $f(z_{it}|z_{it} \leq \delta)$ and $f(z_{it}|z_{it} > \delta)$, for given values of $\delta \in Q_z$ searched for a threshold, and then using the copula transformation $z_{it}^{*(h)} = \Phi^{-1}(U_{z_{it}^{(h)}})$.

Following the literature on threshold models (see, e.g., Chan (1993), and Tong (2010) for a survey), it can be seen that the estimator of δ obtained by the above procedure is $T-$consistent. The estimates of vector $\theta(\delta)$ which correspond to the optimal estimate of δ, are $\sqrt{T}$ asymptotically normal. Since the elements of variables $z_{it}^{*(h)}$ constitute generated regressors, the efficiency of the above estimator of $\theta(\delta)$ can be further improved by adopting an instrumental variable method, in the second step, as noted before. Alternatively, more accurate estimates of $\theta(\delta)$ can be obtained based on a parametric bootstrap method; this may be proved more appropriately for small

T.

3 Extension to a Simultaneous System of Equations

The copulas method presented before can be extended to a system of N-equations (units). This can be done by exploiting results of copula theory for high-dimension joint distributions. For a small number of units N, often met in macroeconomic studies, the SUR framework constitutes a natural choice to estimate model (1) under endogeneity of the threshold variable. Note that, depending on how large is T compared to N, the SUR framework can allow for heterogeneity in the slope parameters of the model and cross-sectional dependence (see, e.g., Meligotsidou et al. (2014)).[5]

To show how to implement the method in the SUR framework of (1), we assume, for expositional convenience, homogeneity of the slope coefficients of the model (including its intercept) and the threshold parameter δ. Then, using matrix algebra notation, we write the model in a system framework as follows,

$$Y_t = X_t' b^{(1)} \mathcal{I}(z_t \leq \delta) + X_t' b^{(2)} \mathcal{I}(z_t > \delta) + \varepsilon_t, \tag{8}$$

where $Y_t = (y_{1t}, y_{2t}, ..., y_{Nt})'$, $X_t = diag(x_{1t}, x_{2t}, ..., x_{Nt})$, $b^{(1)} = (\beta^{(1)}, \beta^{(1)}, ..., \beta^{(1)})'$, $b^{(2)} = (\beta^{(2)}, \beta^{(2)}, ..., \beta^{(2)})'$, $z_t = (z_{1t}, z_{2t}, ..., z_{Nt})'$ and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{Nt})'$, and where $\mathbb{E}(\varepsilon_t|X_t, z_t) \neq 0$ and $\mathbb{E}(\varepsilon_t|X_t, z_t) = \begin{cases} \mathbb{E}(\varepsilon_t|z_t \leq \delta), & \text{if } z_t \leq \delta \\ \mathbb{E}(\varepsilon_t|z_t > \delta), & \text{if } z_t > \delta \end{cases}$.

Following similar steps to those for the single equation framework, we will substitute the vector of error terms ε_t in system (8) with a linear relationship of the copula-transformed vector of z_t in each regime of the model obtained based on the PITs of the elements of truncated random vectors $z_t^{(1)} = z_t \times \mathcal{I}(z_t \leq \delta)$ and $z_t^{(2)} = z_t \times \mathcal{I}(z_t > \delta)$, respectively. This can be seen more analytically as follows.

Using multivariate copulas, write the joint probability densities $f(\varepsilon_t, z_t|z_t \leq \delta)$ and

[5] If the slope coefficients of the explanatory variables of the model are homogenous across i, we can also employ panel data estimation methods, which can considerably improve the efficiency of the estimation and inference procedures of both the threshold parameter and slope parameters of the model, by exploiting both the time (T) and cross-section (N) dimensions of our data. This can be done in cases where $N > T$.

$f(\varepsilon_t, z_t | z_t > \delta)$ as,

$$f(\varepsilon_t, z_t | z_t^{(1)}) \equiv f(\varepsilon_t, z_t | z_t \leq \delta) = c^{(1)}(U_{\varepsilon_{1t}}, ...U_{\varepsilon_{Nt}}, U_{z_{1t}^{(1)}}, ..., U_{z_{Nt}^{(1)}}) \prod_{i=1}^{N} f(\varepsilon_{it}) \prod_{i=1}^{N} f(z_{it}|z_{it}^{(1)}) \tag{9a}$$

$$f(\varepsilon_t, z_t | z_t^{(2)}) \equiv f(\varepsilon_t, z_t | z_t > \delta) = c^{(2)}(U_{\varepsilon_{1t}}, ...U_{\varepsilon_{Nt}}, U_{z_{1t}^{(2)}}, ..., U_{z_{Nt}^{(2)}}) \prod_{i=1}^{N} f(\varepsilon_{it}) \prod_{i=1}^{N} f(z_{it}|z_{it}^{(2)}) \tag{9b}$$

where $c^{(h)}(U_{\varepsilon_{1t}}, ...U_{\varepsilon_{Nt}}, U_{z_{1t}^{(h)}}, ..., U_{z_{Nt}^{(h)}})$ are multivariate copulas, and $U_{\varepsilon_{it}}$ and $U_{z_{it}^{(h)}}$ are uniform $(0, 1)$ variables, for all i. In particular, $U_{\varepsilon_{it}}$, $U_{z_{it}^{(1)}}$ and $U_{z_{it}^{(2)}}$ constitute the PITs of random variables ε_{it}, $z_{it}^{(1)}$ and $z_{it}^{(2)}$, respectively, for all i. If we assume that z_{it} are independent across i, then the conditional densities $f(\varepsilon_t|z_t^{(h)})$ can be defined based on (9a)–(9b) as follows:

$$f(\varepsilon_t|z_t^{(h)}) = \frac{f(\varepsilon_t, z_t|z_t^{(h)})}{f(z_t^{(h)})} = c^{(h)}(U_{\varepsilon_{1t}}, ...U_{\varepsilon_{Nt}}, U_{z_{1t}^{(h)}}, ..., U_{z_{Nt}^{(h)}}) \prod_{i=1}^{N} f(\varepsilon_{it}), \tag{10}$$

for $h = \{1, 2\}$, where $f(z_t^{(h)}) = \prod_{i=1}^{N} f(z_{it}|z_{it}^{(h)})$ is a joint probability function of truncated random variables $z_{1t}^{(h)}$, $z_{2t}^{(h)}$, $\ldots z_{Nt}^{(h)}$. Analytical forms of $f(\varepsilon_t|z_t^{(h)})$ and, hence, conditional expectation terms $\mathbb{E}(\varepsilon_t|z_t \leq \delta)$ and $\mathbb{E}(\varepsilon_t|z_t > \delta)$ can be derived following similar steps to those in the single equation case (recall **Section 2**) and using the multivariate Gaussian copula, defined as,

$$\begin{aligned} c^{(h)}(U_{\varepsilon_{1t}}, ...U_{\varepsilon_{Nt}}, U_{z_{1t}^{(h)}}, ..., U_{z_{Nt}^{(h)}}) &= \phi^{(h)}(\varepsilon_{1t}^*, ..., \varepsilon_{Nt}^*, z_{1t}^{*(h)}, ..., z_{Nt}^{*(h)}) \\ &= \frac{1}{|R^{(h)}|^{1/2}} \exp\left(-\frac{1}{2} U_t^{*(h)\prime} (R^{(h)-1} - I) U_t^{*(h)}\right), \end{aligned} \tag{11}$$

where $\varepsilon_{it}^* = \Phi^{-1}(U_{\varepsilon_{it}})$ and $z_{it}^{*(h)} = \Phi^{-1}(U_{z_{it}^{(h)}})$, for all i and $h = \{1, 2\}$, $U_t^{*(h)\prime} = (\varepsilon_t^{*\prime} \vdots z_t^{*(h)\prime})$ is an $(N + N) \times 1$ dimension vector stacking standardized random variables ε_{it}^* and $z_{it}^{*(h)}$ across i, and $R^{(h)}$ is the correlation matrix of vector $U_t^{*(h)}$. Let us partition correlation matrix $R^{(h)}$ as[6]

$$R^{(h)} = \begin{bmatrix} R_{\varepsilon^*\varepsilon^*} & R_{\varepsilon^* z^{*(h)}} \\ R'_{z^{*(h)}\varepsilon^*} & R_{z^{*(h)} z^{*(h)}} \end{bmatrix},$$

[6] $R_{\varepsilon^*\varepsilon^*}$ is the correlation matrix of vector ε_t^*, $R_{\varepsilon^* z^{*(h)}}$ (or $R'_{z^{*(h)}\varepsilon^*}$) is the correlation matrix between vectors ε_t^* and $z_t^{*(h)}$, and $R_{z^{*(h)} z^{*(h)}}$ is the correlation matrix of vector $z_t^{*(h)}$.

Then, the closed-form solutions of pdf $f(\varepsilon_t|z_t^{(h)})$ can be derived based on (9a, 9b) under the assumption that the elements of the copula-transformed vector $z_t^{*(h)}$, which corresponds to $z_t^{(h)}$, are independent across i, made above. This assumption means that $R_{z^{*(h)}z^{*(h)}} = I_N$, where I_N is an $(N \times N)$ dimension matrix.

The elements of the vector of error terms ε_t^* (or ε_t) are allowed to be correlated across i, which means that we can estimate model (1) in the SUR based framework (8) using a GLS estimation procedure. Under these assumptions, we can show that the following linear relation holds between ε_t and the vector of copula-transformed variables $z_t^{*(h)}$

$$\varepsilon_t = \Lambda^{(h)} z_t^{*(h)} + \left(\Omega^{(h)}\right)^{1/2} e_t^{(h)}, \tag{12}$$

where $\Lambda^{(h)} = \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*}$, $\Sigma_\varepsilon = diag[\sigma_{\varepsilon_1}^2, \sigma_{\varepsilon_2}^2, ..., \sigma_{\varepsilon_N}^2]$, $e_t^{(h)}$ is an $(N \times 1)$ dimension vector of $NIID(0,1)$ random variables, which is independent of $z_t^{*(h)}$, and
$\Omega^{(h)} = \left(\Sigma_\varepsilon^{1/2} R_{\varepsilon^*\varepsilon^*} \Sigma_\varepsilon^{1/2} - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} \Sigma_\varepsilon^{1/2}\right)$.

Proof: Appendix B

The relationship given by (12) implies that $\mathbb{E}(\varepsilon_t|z_t \leq \delta) = \Lambda^{(1)} z_t^{*(1)}$ and $\mathbb{E}(\varepsilon_t|z_t > \delta)$=$\Lambda^{(2)} z_t^{*(2)}$, which means that we can consistently estimate the threshold parameter δ and the vector of the slope parameters of (8), $b^{(h)}$, based on the following augmented version of the SUR model (8):

$$Y_t = (X_t' b^{(1)} + \Lambda^{(1)} z_t^{*(1)})\mathcal{I}(z_t \leq \delta) + (X_t' b^{(2)} + \Lambda^{(2)} z_t^{*(2)})\mathcal{I}(z_t > \delta) + e_t, \tag{13}$$

where e_t is a vector of error terms which are defined above across the two regimes.

Next, we make the following remarks concerning the estimation and the specification of system (13) and the applicability of our method. First, the estimation of the system can be carried out along the lines of the two-step estimation method of the single-equation model (7), described previously. The estimate of δ will be superconsistent, while those of the remaining parameters of the model will be $\sqrt{T}$ consistent. As noted before, where ε_{it} are correlated across the equations of the system, the estimates of the slope parameters can be obtained based on the GLS estimator.

One problem concerning the estimation and inference procedure of system (13) suggested above may be the lack of degrees of freedom. This will appear when N is high relative to T, as the rhs of the system will include a large number of copula transformed variables. This number of regressors can be reduced considerably, when z_{it} is a common threshold variable (or an index/factor) across all i and/or z_{it} and ε_{it} are uncorrelated for some equations (units i) of the system. That is, $\rho^{(h)} = 0$, for $i \neq j$, which means that $R_{\varepsilon^* z^{*(h)}} = R_{z^{*(h)}\varepsilon^*} = diag[\rho_1^{(h)}, \rho_2^{(h)}, ..., \rho_N^{(h)}]$. This case is considered in our simulation study. Obviously, the actual specification of the threshold variable z_{it} across i is an empirical matter.

Our second remark concerns the applicability of the method when threshold variables z_{it} are not independent across i. Then, we can rely on a linear model of variables z_{it}, e.g., $z_{it} = w_{it}'\pi + u_{it}$ where u_{it} are independent across i. Instead of expectation terms $\mathbb{E}(\varepsilon_t | z_t \leq \delta)$ and $\mathbb{E}(\varepsilon_t | z_t > \delta)$, we can employ copula transformations of $\mathbb{E}(\varepsilon_t | u_t \leq \iota\delta - W_t\Pi)$ and $\mathbb{E}(\varepsilon_t | u_t > \iota\delta - W_t\Pi)$, respectively, where $u_t = (u_{1t}, u_{2t}, ..., u_{Nt})'$, term ι is a vector of ones and W_t is a matrix of the vector of variables w_{it}' for all i (recall our remarks in section 2.1).

4 Monte Carlo

In this section, we evaluate the performance of our suggested method to control for the endogeneity of the threshold variable based on a MC study. The main aim of our analysis is to appraise the small sample performance of the estimation procedure of threshold parameter δ, suggested before, controlling for the endogeneity of the threshold variable. In our MC exercise, we assume that the data generating process (DGP) of the threshold model is given as follows:

$$y_{it} = (\beta_1^{(1)} + \beta_2^{(1)} x_{2it} + \beta_3^{(1)} x_{3it})\mathcal{I}(z_{it} \leq \delta) + (\beta_1^{(2)} + \beta_2^{(2)} x_{2it} + \beta_3^{(2)} x_{3it})\mathcal{I}(z_{it} > \delta) + \varepsilon_{it}, \quad (14)$$

where $x_{2it} \sim NIID(0.25, 1)$, $x_{3it} \sim NIID(0.75, 1)$. We consider different cases of distributions for the error term ε_{it} and threshold variable z_{it}. The error structure and the distributions of ε_{it} and z_{it} considered in our simulation exercise are given in **Table 1** – **Table 2**.

In the exercise, we assume that,

$$\beta_1^{(1)} = \beta_3^{(1)} = 1.0, \;\; \beta_2^{(1)} = 2.0$$

$$\beta_1^{(2)} = \beta_3^{(2)} = 0.0, \;\; \beta_2^{(2)} = 1.0$$

and we present results for the cases that the correlation coefficients between ε_{it} and z_{it} are given by $\rho = \{0.00, 0.55, 0.75\}$.[7] For all different cases of the distributions of ε_{it} and z_{it} considered in our analysis, we set the threshold value to the 75% percentile of the distribution of z_{it}, for all i. We also examined threshold values at the 25% percentile of the distribution of z_{it} (often also considered in the literature) but the results do not change the main conclusions of our MC study. We consider sample sizes of $T = \{100, 300\}$, and cases of $N = \{1, 5, 10\}$, where $N = 1$ corresponds to the single regression framework of (1).

We carry out $1,000$ MC iterations. In each iteration, we estimate the threshold parameter δ as suggested in the previous section based on augmented regression model (7), for the single equation case, and on (13), for the SUR case. In **Table 1** – **Table 2**, we present the average values of the bias of the estimates of δ from its true value and the root mean square error (RMSE) of these estimates, over all iterations. The first table presents estimates for the case that ε_{it} and z_{it} are both normally distributed, while the second table assumes that they are Student-t distributed.

In these two tables, we also present estimates of the above metrics of δ for the case that we ignore the endogeneity of z_{it}, by estimating the model without the regressors adjusting for the endogeneity of the threshold variable, and for the case that we estimate δ based on the Kourtellos et al. (2016) approach, which adjusts (14) for the endogeneity of the threshold variable z_{it} using the inverse Mills ratio terms. The latter estimates are given in '$\langle\rangle$' parentheses. These extra sets of estimates indicate how serious is the problem of threshold variable endogeneity if it is ignored, in practice, and in the case that the distribution of the threshold variable is not normally distributed.

Results suggest the following conclusions. First, ignoring the endogeneity of threshold variable z_{it} causes serious biases in the estimates of threshold parameter δ. It tends to underestimate the true value of δ, substantially. As expected, the bias is greater in magnitude in the case that the correlation between z_{it} and ε_{it}, ρ, is high and/or the size of T is small, compared to the cases that ρ is small and/or T is large. These results hold for both the single regression and SUR based estimates of model (14), ignoring the endogeneity of z_{it}.

[7] Note that the values of the parameters of the model considered in our simulation analysis are close to those considered in the simulation studies of Lundbergh et al. (2003), Caner and Hansen (2004), and Kourtellos et al. (2016).

Second, our method can successfully control for the endogeneity problem of the threshold variable on the estimates of threshold parameter δ. It reduces the estimation bias of δ, substantially, even for small T (i.e., $T = 100$) and/or very high values of ρ (i.e., ρ =0.75). As expected, the reduction of the magnitude of the bias is bigger as both T and N increase, i.e., $T = 300$ and $N = 10$, compared to $T = 100$ and $N = 5$. The bias reduces fastest with respect to the T dimension, compared to the N. An analogous conclusion can be drawn for the RMSE. This is found to be quite substantial for small values of T (i.e., $T = 100$) due to weak threshold effects assumed in our MC study (see also Gonzalo and Pitarakis (2002)). Note that, for the case that both ε_{it} and z_{it} are normally distributed, the performance of our method successfully compares to that of Kourtellos et al. (2016), (as might be expected).

However, for the case that both ε_{it} and z_{it} are Student-t distributed, our method clearly outperforms that of Kourtellos et al., which relies on the assumption that the threshold variable is normally distributed.[8] These results indicate that deviations from the normality assumption of the threshold variable may also cause serious biases in the estimates of δ, and thus motivate the use of our method based on copula theory to eliminate these biases.

Finally, another useful conclusion that can be drawn from the tables concerns the case of no endogeneity of z_{it} (i.e., $\rho = 0$). For this case, the results of the tables indicate that estimation of the augmented models (7) and (13) lead to unbiased estimates of δ. This is true even for small T. This result means that estimating the augmented model to control for threshold variable endogeneity effects on it does not influence the estimates of threshold parameter δ, when there is no such effects. Thus, if there are no serious degrees of freedom problems, one can estimate the above models, directly, and then test if there are significant threshold effects within them. This is a useful exercise in cases where the information criteria fail to detect linearity against threshold effects, as shown below.

[8] Similar results are found for other distributions of the threshold variable z_{it}, like the gamma distribution and a mixture of normals distributions. These results are not reported for reasons of space.

TABLE 1: MC: THRESHOLD PARAMETER δ (NORMALLY DISTRIBUTED ERRORS),
$\varepsilon_{it} = v_{it} + \eta_{it}; z_{it} = \delta + c_i v_{it} + \zeta_{it}; \delta = 3.9; v_{it}, \eta_{it}, \zeta_{it} \sim NIID(0,1)$

	Ignoring the Endogeneity of z_{it}				Controlling for the Endogeneity of z_{it}							
	$T = 100$		$T = 300$		$T = 100$				$T = 300$			
ρ	Bias	RMSE	Bias	RMSE	Bias		RMSE		Bias		RMSE	
	N = 1											
0.00	-0.002	0.368	-0.026	0.039	-0.137	⟨-0.35⟩	0.281	⟨0.53⟩	-0.034	⟨-0.04⟩	0.049	⟨0.06⟩
0.55	-0.576	0.927	-0.291	0.344	-0.310	⟨-0.24⟩	0.476	⟨0.47⟩	-0.096	⟨-0.01⟩	0.083	⟨0.04⟩
0.75	-0.637	1.039	-0.356	0.436	-0.351	⟨-0.18⟩	0.485	⟨0.38⟩	-0.082	⟨-0.00⟩	0.075	⟨0.03⟩
	N = 5											
0.00	-0.003	0.188	0.005	0.002	-0.076	⟨-0.05⟩	0.615	⟨1.02⟩	-0.002	⟨0.004⟩	0.003	⟨0.005⟩
0.55	-0.425	1.202	-0.244	0.350	0.035	⟨-0.00⟩	0.667	⟨0.58⟩	-0.003	⟨0.007⟩	0.009	⟨0.03⟩
0.75	-0.681	1.338	-0.502	0.723	-0.111	⟨0.19⟩	0.307	⟨1.11⟩	-0.009	⟨0.26⟩	0.008	⟨1.22⟩
	N = 10											
0.00	0.010	0.095	0.0017	0.0005	-0.031	⟨-0.008⟩	0.654	⟨1.14⟩	-0.0008	⟨0.01⟩	0.0006	⟨0.02⟩
0.55	-0.450	1.810	-0.271	0.320	-0.067	⟨0.09⟩	1.26	⟨0.85⟩	-0.0007	⟨0.003⟩	0.009	⟨0.002⟩
0.75	-0.762	1.394	-0.576	0.819	-0.060	⟨0.52⟩	0.272	⟨2.19⟩	-0.017	⟨0.450⟩	0.008	⟨2.15⟩

Notes: The table presents average values of the bias and RMSE of the estimates of threshold parameter δ of model (1). This is done for the case ignoring the endogeneity of the threshold variable z_{it} and that controlling for it, by estimating the augmented models (7) and (13). In our simulation exercise, we assume that $\beta_1^{(1)} = \beta_3^{(1)} = 1.0$, $\beta_2^{(1)} = 2.0$, and $\beta_1^{(2)} = \beta_3^{(2)} = 0.0$, $\beta_2^{(2)} = 1.0$ and we choose values of coefficients c_i which imply the following values of the correlation coefficients between ε_{it} and z_{it}: $\rho = \{0.00, 0.55, 0.75\}$. The table considers the case that the error term ε_{it} is normally distributed. Note that in '⟨⟩' parentheses, we report values of the bias and RMSE metrics of δ based on Kourtellos' et al. (2016) method.

TABLE 2: MC: THRESHOLD PARAMETER δ (STUDENT-T ERRORS),
$\varepsilon_{it} = v_{it}; z_{it} = \delta + c_i v_{it} + \zeta_{it}; \delta = 3.9; v_{it}, \zeta_{it} \sim t_{DF=5}$

	Ignoring the Endogeneity of z_{it}				Controlling for the Endogeneity of z_{it}							
	$T = 100$		$T = 300$		$T = 100$				$T = 300$			
ρ	Bias	RMSE	Bias	RMSE	Bias		RMSE		Bias		RMSE	
	N = 1											
0.00	0.026	2.257	0.045	0.338	0.260	$\langle 0.948 \rangle$	3.511	$\langle 7.159 \rangle$	0.210	$\langle 2.321 \rangle$	0.973	$\langle 14.435 \rangle$
0.55	-0.604	2.389	-0.394	1.036	-0.218	$\langle -1.732 \rangle$	0.872	$\langle 10.035 \rangle$	0.131	$\langle 3.352 \rangle$	0.204	$\langle 20.515 \rangle$
0.75	-0.655	2.417	-0.455	1.075	-0.205	$\langle 3.105 \rangle$	0.293	$\langle 18.570 \rangle$	-0.067	$\langle 3.106 \rangle$	0.128	$\langle 17.010 \rangle$
	N = 5											
0.00	-0.004	0.004	0.051	0.002	-0.005	$\langle 0.572 \rangle$	0.005	$\langle 1.031 \rangle$	0.002	$\langle 0.414 \rangle$	0.001	$\langle 1.858 \rangle$
0.55	-0.609	1.614	-0.450	0.570	-0.016	$\langle 0.694 \rangle$	0.104	$\langle 2.468 \rangle$	0.004	$\langle 0.614 \rangle$	0.002	$\langle 2.471 \rangle$
0.75	-1.081	2.322	-1.182	2.201	-0.009	$\langle 1.172 \rangle$	0.007	$\langle 4.237 \rangle$	-0.016	$\langle 2.415 \rangle$	0.010	$\langle 11.587 \rangle$
	N = 10											
0.00	-0.001	0.002	0.000	0.001	-0.004	$\langle 0.505 \rangle$	0.002	$\langle 1.253 \rangle$	0.000	$\langle 1.478 \rangle$	0.000	$\langle 8.092 \rangle$
0.55	-0.625	0.892	-0.425	0.399	-0.012	$\langle 1.137 \rangle$	0.003	$\langle 3.978 \rangle$	0.007	$\langle 3.294 \rangle$	0.001	$\langle 12.920 \rangle$
0.75	-1.127	1.796	-1.135	1.670	0.000	$\langle 2.126 \rangle$	0.002	$\langle 7.278 \rangle$	0.005	$\langle 4.769 \rangle$	0.001	$\langle 29.989 \rangle$

Notes: The table presents average values of the bias and RMSE of the estimates of threshold parameter δ of model (1). This is done for the case ignoring the endogeneity of the threshold variable z_{it} and that controlling for it, by estimating the augmented models (7) and (13). In our simulation exercise, we assume that $\beta_1^{(1)} = \beta_3^{(1)} = 1$, $\beta_2^{(1)} = 2$, and $\beta_1^{(2)} = \beta_3^{(2)} = 0$, $\beta_2^{(2)} = 1.0$ and we choose values of coefficients c_i which imply the following values of the correlation coefficients between ε_{it} and z_{it}: $\rho = \{0.00, 0.55, 0.75\}$. The table considers the case that the error term ε_{it} is Student-t distributed. Note that in '⟨⟩' parentheses, we report values of the bias and RMSE metrics of δ based on Kourtellos' et al. (2016) method.

5 Estimation of a Threshold Foreign Trade Multiplier

We now implement the method to estimate a threshold model of the foreign trade multiplier net of possible threshold variable endogeneity effects. This multiplier is a standard tool to quantify the effects of exports and on real economic growth (see, e.g., Robinson (1952), and Lego et al. (2000)). A value above (below) unity means that increasing exports as a percentage of real output will increase output more (less) than the initial export increase. We consider a threshold model of this multiplier, following recent evidence that exports critically depend on regime shifts in the real exchange rate and its volatility. These shifts can capture differences between flexible and stable (or pegging) exchange rate regimes (see, e.g., Tenreyro (2007), Berman and Berthou (2009), Berman and Martin (2012), and Aghion et al. (2009)). Almost certainly, these shifts will affect the estimates of the foreign trade multiplier.

There are two main views on how exchange rate volatility affects exports and, hence, growth. The first claims that this volatility has negative effect on exports, since it reduces international trade and discourages investment (see, e.g., Grier and Smallwood (2013)). Alternatively, some theories argue that exchange rate volatility has positive effect on both trade and growth, since it absorbs institutional and macroeconomic differences across countries (see, e.g., Yeyati and Sturzenegger (2003), and Edwards and Yeyati (2005)). Supporters of this view argue that persistency of exchange rate volatility encourage exporters to increase their export volumes and supply in order to maximize their profits and/or revenues (see Franke (1991), Kroner and Lastrapes (1993), and Egert and Morales-Zumaquero (2008)).

To assess whether the foreign trade multiplier relationship depends on the regime of exchange rate volatility, we estimate the following threshold model:[9]

$$\tilde{y}_{it} = a_{it} + \beta^{(1)}\tilde{x}_{it}\mathcal{I}(z_{it} \leq \delta) + \beta^{(2)}\tilde{x}_{it}\mathcal{I}(z_{it} > \delta) + \varepsilon_{it}, \tag{15}$$

where $\tilde{y}_{it} = \frac{y_{it}-y_{it-1}}{y_{it-1}}$ and $\tilde{x}_{it} = \frac{x_{it}-x_{it-1}}{y_{it-1}}$, where y_{it} stands for the real GDP at 2010 reference levels, x_{it} denotes exports of goods and services in 2010 prices, z_{it} is a measure of the real effective exchange rate volatility (denoted as V^{REER}), which plays the role of the threshold variable, and ε_{it} is the regression error term. The real effective exchange rate ($REER$) measures the real value of country's i currency against the basket of its trading partners. Thus, it can capture the relative level of competitiveness of country i against its trading partners, the drivers of trade flows and the long-run equilibrium value of a currency. As also argued by Baggella et al. (2006), compared to bilateral real exchange rate, $REER$ has following interesting features. First, it can better capture regional integration (trade partner's) effects on trade flows and growth. Second, it can

[9]Note that this relationship corresponds to that of Kraay (2012), without threshold effects, used to obtain estimates of the government spending multiplier.

smooth out favorable and unfavorable exchange rate movements with trade partners and, third, it is less sensitive to the bilateral exchange rate volatility with the US dollar.

In our empirical analysis, REER is calculated from the nominal effective exchange rate ($NEER$) and a measure of the relative price of country i and its trading partners j, for all $j \neq i$ (see e.g., Darvas (2012)): $REER_{it} = \frac{NEER_{it} \times CPI_{it}}{CPI^{*}_{it}}$, where $NEER_{it} = \prod_{j \neq i}^{N} s_{it}(j)^{w_{ij}}$ is the geometric mean of the nominal bilateral exchange rates of a country i with its trading partners j, w_{ij} is the weight of trading partner j, CPI_{it} is the consumer price index of country i, and $CPI^{*}_{it} = \prod_{j \neq i}^{N} CPI_{j}^{w_{ij}}$ is the geometric mean of the CPI's of trading partners j. Following Frankel and Wei (1993) and, more recently, Tenreyro (2007), the volatility of $REER_{it}$ (denoted V_{it}^{REER}), which plays the role of threshold variable z_{it} in our model, is calculated as the standard deviation of the first difference of the logarithms of $REER_{it}$, over a rolling window of five years observations.[10] As aptly noted by Tenreyro (2007), exchange rate volatility is expected to be endogenous to exports and trade, as well as real economic growth shocks.

[10]We also considered alternative estimates of V_{it}^{REER} based on a GARCH process, but we found that these do not change the results of our analysis.

TABLE 3: DESCRIPTIVE STATISTICS

	Canada			France			Italy			Japan			UK			US			Netherlands		
	$\tilde{y}$	$\tilde{x}$	V_{it}^{REER}	$\tilde{y}$	$\tilde{x}$	V_{it}^{REER}	$\tilde{y}$	$\tilde{x}$	V_{it}^{REER}	$\tilde{y}$	$\tilde{x}$	V_{it}^{REER}	$\tilde{y}$	$\tilde{x}$	V_{it}^{REER}	$\tilde{y}$	$\tilde{x}$	V_{it}^{REER}	$\tilde{y}$	$\tilde{x}$	V_{it}^{REER}
Mean	0.031	0.010	0.037	0.026	0.008	0.0344	0.022	0.007	0.038	0.035	0.005	0.082	0.023	0.008	0.052	0.029	0.004	0.037	0.026	0.023	0.033
Std. Dev.	0.021	0.014	0.016	0.019	0.008	0.014	0.026	0.011	0.020	0.036	0.008	0.035	0.021	0.008	0.020	0.021	0.004	0.001	0.020	0.020	0.015
Skewness	-0.921	-0.923	-0.566	-0.087	-1.483	0.785	-0.217	-2.37	0.850	0.537	-1.757	-0.556	-0.973	-0.517	-0.050	-0.537	-0.787	0.30	-0.560	-0.946	0.158
Excess Kurtosis	1.765	2.682	-0.705	0.319	8.069	0.263	0.670	11.02	0.465	0.536	12.249	-0.296	1.577	4.053	-0.216	0.317	1.818	0.207	0.750	4.818	-1.039
Normality†	–	–	[0.030]	–	–	[0.066]	–	–	[0.332]	–	–	[0.127]	–	–	[0.931]	–	–	[0.053]	–	–	[0.242]
Normality‡	–	–	[0.006]	–	–	[0.070]	–	–	[0.070]	–	–	[0.032]	–	–	[0.782]	–	–	[0.035]	–	–	[0.109]
min	-0.032	-0.041	0.003	-0.029	-0.030	0.012	-0.054	-0.047	0.009	-0.054	-0.036	0.010	-0.043	-0.025	0.009	-0.027	-0.010	0.009	-0.037	-0.062	0.008
max	0.069	0.036	0.060	0.071	0.0289	0.069	0.072	0.027	0.093	0.119	0.031	0.149	0.065	0.033	0.095	0.073	0.013	0.078	0.064	0.069	0.06

Notes: This table presents descriptive statistics: mean, Standard Deviation (Std. Dev.), the skewness and kurtosis coefficients, and the min and max values of all variables employed in the estimation of (15), for all countries i. † and ‡ indicates the Doornik and Hansen (2008) multivariate and Shapiro-Wilks (Shapiro (1980)) Normality tests, respectively, (numbers in squared parentheses represent probability values).

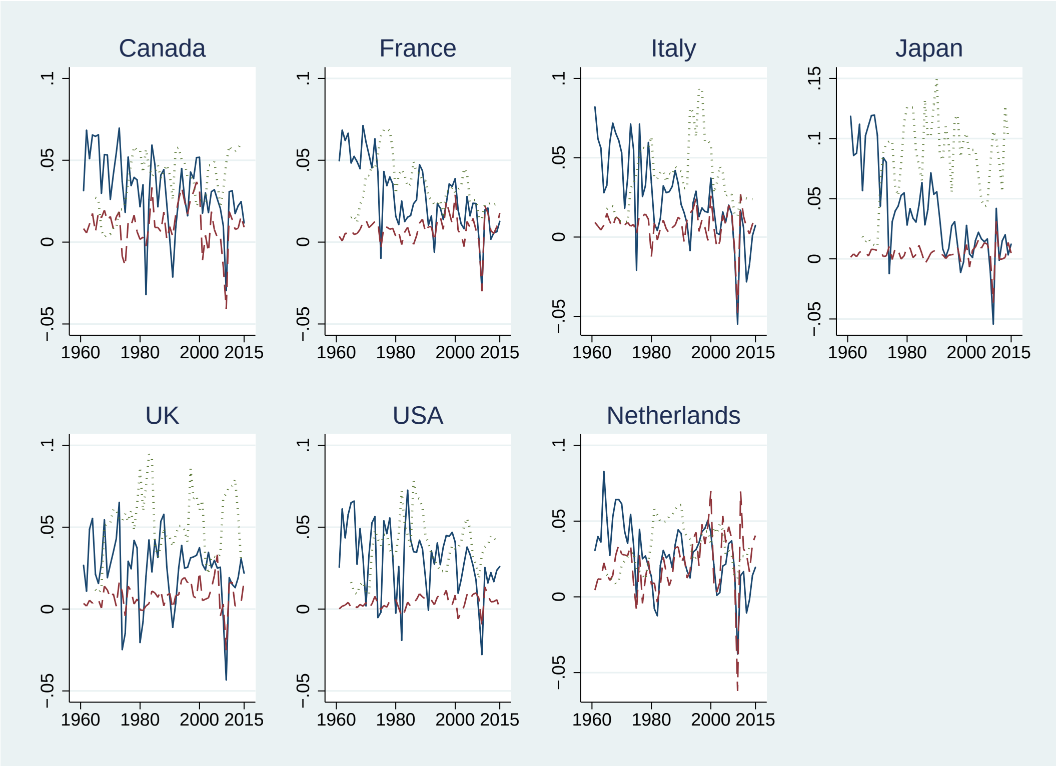
To estimate model (15), we rely on data on the following seven developed countries: Canada, France, Italy, Japan, United Kingdom (UK), United States (US) and Netherlands. Our data covers the period from 1963 to 2015. Descriptive statistics and normality tests for the series employed in the estimation of the model are given in **Table 3**.

Figure 1 graphs the series. Both the real economic growth rate and real exchange volatility series exhibit higher degree of volatility than the explanatory variable of the model $\tilde{x}_{it}$, with the exception of Netherlands. The highest degree of volatility of $\tilde{x}_{it}$ is observed in the years of oil crises (e.g., 1973-1974) and the recent years 2008 and 2009, associated with the global financial crisis. Similar patterns are also observed for the threshold variable V_{it}^{REER}, with the exception of Japan and the UK. For these two countries, V_{it}^{REER} seems to fluctuate considerably over the whole sample, especially for Japan. Another interesting comment that can be made by the inspection of the graphs is that there exists a clear-cut positive relationship between $\tilde{x}_{it}$ and V_{it}^{REER} after the first oil price crisis and in the middle of nineties, for most countries examined. This is also confirmed by estimates of the correlation coefficients between these two variables, which are not reported in the table for reasons of space.

In our empirical analysis, we estimate model (15) without and with threshold effects. The model without threshold effects constitutes a linear specification of it, assuming $\beta = \beta^{(1)} = \beta^{(2)}$. The estimates of the linear case (**Panel A**) together with the threshold model which ignores the endogeneity of the threshold variable (**Panel B**) are given in **Table 4**.

Table 5 presents the estimates of the threshold model which controls for the endogeneity of the threshold model. Note that, for comparison reasons, the table presents estimates based on our suggested methodology, presented in the previous sections (**Panel B**), and that of Kourtellos et al (2016) (**Panel A**). To investigate if the threshold model (15) constitutes a better specification of the data compared to its linear specification (without threshold effects), in **Table 5** we present values of information criteria AIC and BIC. To formally examine if they are significant differences in slope coefficients $\beta^{(h)}$ between the two regimes of the model, we present estimates of the

FIGURE 1: GROWTH, EXPORTS AND REAL EXCHANGE RATE VOLATILITY



Notes: Solid: Real GDP Growth, Dashed: Exports, Dotted: Real Exchange Rate Volatility.

following likelihood ratio (LR) test statistic:

$$\sup\text{-}LR \equiv \sup_{\delta \in Q_z} \left(-2\left(\log L(\theta(\delta)) - \log L(\beta)\right)\right), \tag{16}$$

where $\log L(\cdot)$ constitutes the maximum log-likelihood function of model (15) under the following null and alternative hypotheses: H_0: $\beta^{(1)} = \beta^{(2)} = \beta$ (no threshold effects) and H_a: $\beta^{(1)} \neq \beta^{(2)}$, respectively. The vectors of the slope parameters of the model under the above alternative and hypotheses are defined as $\theta(\delta) = (\beta^{(1)}, \beta^{(2)}, \lambda^{(1)}, \lambda^{(2)})'$ and $\beta = \beta^{(1)} = \beta^{(2)}$, respectively. Rejection of the above H_0 based on test statistic $\sup\text{-}LR$ means acceptance of its alternative H_a, which provides support of threshold model (15). This is critical in inferring if model (15) constitutes a better specification of the data, compared to its linear counterpart. It can be accompanied with the information criteria AIC and BIC.

Since the threshold and slope parameters are not identified under the above null hypothesis (see, e.g., Hansen (1999)), the distribution of statistic $\sup\text{-}LR$ is not standard. Its critical value (or the probability of type I error) of rejecting null hypothesis

H_0: $\beta^{(1)} = \beta^{(2)}$, at a given level of significance, can be obtained based on a bootstrap statistical method (see, e.,g., Chan (1993) and Hansen (1996)). Below, we present the steps of this method for the single equation based framework.

Step 1. Estimate model (15) under null H_0: $\beta^{(1)} = \beta^{(2)}$, save the slope parameter estimates and calculate the residuals, denoted $\hat{\varepsilon}_{it}$

Step 2. For each bootstrap sample $b = 1, 2, ..., B$, draw a random sample of $t = 1, 2, ..., T$ from the distribution of $\hat{\varepsilon}_{it}$ and calculate the wild bootstrap residuals defined as $\hat{\varepsilon}_{it}^{(b)} = \hat{\varepsilon}_{it}(1 - \eta_{it})$, where η_{it} is a zero-mean and unit-variance random variable generated by Rademacher's distribution, i.e.,

$$\eta_{it} = \begin{cases} 1 & \text{with probability } 0.5 \\ -1 & \text{with probability } 0.5 \end{cases}$$

Step 3. Based on the sample values of $\tilde{x}_{it}$ and z_{it}, the estimates of β under H_0 and bootstrap residuals $\hat{\varepsilon}_{it}^{(b)}$, calculate a bootstrap sample of the dependent variable, denoted $g_{it}^{(b)}$.

Step 4. Given $g_{it}^{(b)}$, estimate model (15) and calculate test statistic $\sup\text{-}LR$, for each bootstrap sample b.

Step 5. Repeat the above 2-4 steps $b = 1, 2, ..., B = 1000$ times. The $(1 - a)\%$ percentile of the empirical distribution of $\sup\text{-}LR$ gives the $a\%$ bootstrap critical value of this statistic. If the computed values of $\sup\text{-}LR$ based on our sample data exceeds this critical value, then we will reject H_0: $\beta^{(1)} = \beta^{(2)}$ against H_a: $\beta^{(1)} \neq \beta^{(2)}$. In **Table 5**, we present the computed values of $\sup\text{-}LR$ based on our sample estimates and their p-values of rejecting the above null hypothesis obtained by the above bootstrap method.

A number of interesting conclusions can be drawn from the results of **Table 4** and **Table 5**. First, they clearly indicate that the threshold model (15) constitutes a better specification of the data compared to the linear model, without threshold effects. This is true for most of the countries examined. This result can be justified by the values of information criteria AIC and BIC, as well as the values of test statistic $\sup\text{-}LR$,reported in the table. There is only one country where $\sup\text{-}LR$ can not reject null

Table 4: Single equation based estimates of (15)

	Canada	France	Italy	Japan	UK	US	Neth.
	Panel A: Linear Specification						
β	0.89 (0.17)	1.20 (0.30)	1.09 (0.29)	1.23 (0.59)	1.14 (0.30)	1.41 (0.67)	0.55 (0.12)
AIC	-130.86	-131.26	-115.42	-95.76	-126.86	-122.02	-129.79
BIC	-128.57	-129.35	-113.50	-93.85	-124.95	-120.11	-127.88
	Panel B: Ignoring Endogeneity of the threshold variable						
$\beta^{(1)}$	1.10 (0.20)	0.98 (0.26)	1.92 (0.38)	13.05 (1.90)	0.80 (0.37)	0.45 (0.33)	1.15 (0.16)
$\beta^{(2)}$	0.40 (0.30)	2.71 (0.46)	0.28 (0.38)	1.02 (0.43)	1.51 (0.38)	3.07 (1.10)	0.31 (0.11)
δ	0.048	0.035	0.020	0.016	0.054	0.045	0.027
AIC	-131.29	-137.32	-119.85	-110.30	-127.09	-122.83	-138.74
BIC	-128.52	-134.45	-116.31	-107.51	-124.22	-119.96	-135.87

Notes: This table presents the single equation for the linear version of the model (without threshold effects, assuming $\beta = \beta^{(1)} = \beta^{(2)}$) and for the full specification of it which assumes $\beta^{(1)} \neq \beta^{(2)}$) (but without endogeneity of the threshold assumed). Numbers in parentheses represent standard errors. AIC (Akaike) and BIC (known as Bayesian, or Schwartz) are information criteria.

hypothesis H_0: $\beta^{(1)} = \beta^{(2)}$. This is Canada, where $\hat{\beta}^{(1)}$ and $\hat{\beta}^{(2)}$ are found to be very close across the high and low real exchange volatility regimes identified throughout the model.

The values of the information criteria AIC and BIC reported in the table indicate that the version of the model controlling for threshold variable endogeneity effects constitutes a much better specification of the data than that of ignoring them. This is more evident for the following countries: Japan, UK, US and Netherlands. For these countries, the estimates of the multiplier coefficients $\beta^{(1)}$ and $\beta^{(2)}$ differ substantially between the estimates of the threshold model ignoring and controlling for the endogeneity of the threshold variable. This is true for both the estimates of the threshold model based on our approach to dealing with the endogeneity of the threshold variable and those based on Kourtellos et al. The differences in the parameter estimates of the threshold model between these two approaches are more apparent for the countries that the normality assumption of the threshold variable is clearly rejected. This mainly happens for countries France, Japan and US where the normality assumption of the threshold variable can be rejected by the data, at the 5% significance level, according to the tests reported in **Table 3** (or marginally above 5% as in the case of France.)

Another interesting conclusion that can be drawn from **Table 5** concerns the estimates of multiplier coefficients $\beta^{(1)}$ and $\beta^{(2)}$ themselves. For the four out of the seven countries examined (namely, France, Japan, UK and US), the estimates of these coefficients are higher in the high exchange rate volatility regime (defined as "2"), compared to the low one (defined as "1"). Actually, the estimates of $\beta^{(2)}$ for the above group of countries exceed unity. The opposite result holds for Italy and the Netherlands. For these two countries, our results indicate that the effects of trade on growth become larger under the low REER volatility regime. For Canada, both $\hat{\beta}^{(1)}$ and $\hat{\beta}^{(2)}$ are below unity and do not differ significantly from each other, as noted before. For the US, Japan and UK, note that, in the low REER volatility regime, there do not exist significant foreign trade multiplier effects on real economic growth rate, as the estimates of $\beta^{(1)}$ are not significantly different from zero, at 5%. Finally, for the countries for which the foreign trade multiplier is not significant in the low-volatility regime, we also present estimates imposing the restriction $\beta^1 = 0$. The results of these estimates

(**Panel B'**) do not change the remaining coefficient estimates of the model, including the threshold parameter; these coefficient estimates remain robust.

Concerning the estimates of threshold parameter δ, the results of the table show significant differences among the countries examined. The country with the highest value of δ is found to be Japan, while that with the smallest one is Italy. For both these countries, note that the estimate of the threshold parameter corresponds to a percentile of the empirical distribution of V^{REER} which is less than 50%. The rest of the countries have values of δ close to each other, which corresponds to higher percentiles of the distribution of V^{REER}. For all the above cases, note that the estimates of δ reported in the table leave a sufficient number of observations in each regime, meaning that our model can identify the two regimes of the exchange rate volatility from the data sufficiently.

Summing up, the results of our empirical estimates indicate that, for most of the G7 countries considered, the foreign trade multiplier effects on economic growth depends on the specific real effective exchange rate volatility regime. For countries France, UK, Japan, UK and US, these multiplier effects become bigger in magnitude at the high exchange rate volatility regime. These results adds support to the view that exchange rate volatility may favor exports and economic growth. They are also consistent with the fact that a very big proportion of the export volumes of the above G7 countries is in machinery, electronic equipment, vehicles and medicals, whose industries are revenue maximizes (see the discussion in Egert and Morales-Zumaquero (2008)).

5.1 The SUR Case Allowing for Slope Heterogeneity

To see if our results remain robust in the SUR framework, we estimate the threshold foreign trade multiplier model (15) based on this framework. In **Table 6** we present estimates of model (15) controlling for threshold variable endogeneity effects which are simultaneously obtained for all the countries of our sample, based on the SUR framework. This framework allows for cross-equation correlation of the error terms of the SUR framework, as noted before. The table presents three different categories of estimation results. The first allows for the estimates of the slope and threshold parameters of the model (i.e., $\beta^{(1)}$, $\beta^{(2)}$ and δ) to be heterogenous across all countries i. This set corresponds to that of the single equation estimates reported in **Table 4**.

TABLE 5: SINGLE EQUATION BASED ESTIMATES OF (15)

	Canada	France	Italy	Japan	UK	US	Neth.
Panel A: Controlling for Threshold Endogeneity (Mills Ratio Method)							
$\beta^{(1)}$	0.72	1.83	1.86	-10.13	0.06	1.22	1.06
	(0.29)	(0.46)	(0.38)	(2.99)	(0.48)	(0.70)	(0.16)
$\beta^{(2)}$	0.94	1.47	0.54	1.18	1.75	2.55	0.51
	(0.19)	(0.37)	(0.41)	(0.48)	(0.37)	(1.28)	(0.15)
δ	0.03	0.02	0.02	0.05	0.05	0.05	0.02
AIC	-134.67	-135.85	-119.28	-106.35	-128.24	-125.18	-138.99
$SBIC$	-129.89	-131.07	-114.51	-101.58	-123.46	-120.41	-134.21
Panel B: Controlling for Threshold Endogeneity (Copula Method)							
$\beta^{(1)}$	0.79	1.04	1.86	-0.14	0.04	-1.60	1.07
	(0.22)	(0.29)	(0.38)	(0.62)	(0.47)	(1.49)	(0.16)
$\beta^{(2)}$	0.98	2.62	0.59	1.50	1.74	2.20	0.51
	(0.26)	(0.63)	(0.41)	(0.60)	(0.37)	(0.91)	(0.14)
δ	0.048	0.036	0.019	0.104	0.049	0.033	0.024
Panel B$'$: Controlling for Threshold Endogeneity (Copula Method), Restricted Estimates							
$\beta^{(1)}$				0.00	0.00	0.00	
$\beta^{(2)}$				1.40	1.74	2.19	
				(0.40)	(0.36)	(0.75)	
δ				0.017	0.049	0.033	
# obs in "2"	32.65%	40.82%	81.63	24.48%	61.24%	69.38%	73.47%
sup-LR	0.37	5.52	4.26	4.04	5.30	4.36	5.89
	[0.81]	[0.03]	[0.008]	[0.03]	[0.05]	[0.04]	[0.05]
AIC	-132.34	-135.48	-119.04	-113.49	-129.29	-122.89	-139.40
BIC	-127.56	-130.70	-114.26	-109.66	-125.45	-118.32	-134.62

Notes: This table presents the single equation for the linear version of the model (without threshold effects, assuming $\beta = \beta^{(1)} = \beta^{(2)}$) and for the full specification of it which assumes $\beta^{(1)} \neq \beta^{(2)}$). Standard errors in parentheses. For the full specification of the model, we consider estimates ignoring threshold endogeneity bias effects on the parameter estimates and when we control for these effects. For the specification controlling for the threshold endogeneity effects, we also present restricted estimates of the model in the case where the slope coefficients are not different from zero at the 5% level. %obs in "2" gives the number of observations in regime "2" identified by estimated threshold parameter δ. sup-LR is statistic (16) testing null hypothesis H_0: $\beta^{(1)} = \beta^{(2)}$ against its alternative H_a: $\beta^{(1)} \neq \beta^{(2)}$. p-values of sup-$LR$ obtained by the bootstrap method described in the text are given in squared parentheses.

Table 6: System equation based estimates of (15) controlling for Threshold Endogeneity

	Canada	France	Italy	Japan	UK	US	Neth.	
	Heterogenous threshold and slope parameters							
$\beta^{(1)}$	0.54	0.783	0.85	-0.104	0.21	-0.51	0.40	
	(0.20)	(0.26)	(0.31)	(0.65)	(0.37)	(0.43)	(0.11)	
$\beta^{(2)}$	0.54	1.495	0.56	1.11	1.25	1.48	0.58	
	(0.20)	(0.46)	(0.30)	(0.36)	(0.26)	(0.57)	(0.14)	
δ	NA	0.034	0.020	0.10	0.037	0.030	0.030	
AIC	-977.47							
BIC	-915.74							
	Homogenous threshold and slope parameters							
$\beta^{(1)}$								0.35
								(0.06)
$\beta^{(2)}$								0.72
								(0.14)
δ								0.054
AIC								-970.47
BIC								-949.50
	Limited homogeneity of the threshold and slope parameters							
$\beta^{(1)}$	0.71	0.75	0.80	0.40	0.40	0.40	0.38	
	(0.22)	(0.22)	(0.30)	(0.30)	(0.30)	(0.30)	(0.11)	
$\beta^{(2)}$	0.71	1.20	0.48	1.20	1.20	1.20	0.65	
	(0.22)	(0.21)	(0.30)	(0.21)	(0.21)	(0.21)	(0.16)	
δ	NA	0.035	0.019	0.10	0.035	0.035	0.035	
AIC	-985.67							
BIC	-933.58							

Notes: The table presents SUR (seemingly unrelated regression) based estimates of model (15) controlling for threshold variable endogeneity effects on the parameter estimates of the model, for all countries. The table gives three different sets of results. The first set assumes full heterogeneity of the threshold and slope parameters of the model across countries, while the second assumes homogeneity of these parameters except of the intercepts. The third set considers a limited degree of homogeneity on both the threshold and slope parameters of the model. In particular, it assumes that δ is common for US, UK, Netherlands and France, $\beta^{(2)}$ for US, UK, Japan and $\beta^{(1)}$ for US, UK and Japan. For Canada, it is assumed that $\beta^{(1)} = \beta^{(2)} = \beta$. Standard errors are in parenthesis. AIC and BIC are the Akaike and Bayesian (or Schwartz) information criteria.

The second assumes that $\beta^{(1)}$, $\beta^{(2)}$ and δ are homogenous, for all countries i, while the third set considers a limited degree of homogeneity of $\beta^{(1)}$, $\beta^{(2)}$ and δ across i. In particular, the third set of results assumes that only some of these parameters are the same across i. Given our single equation results, we assume that δ is common for US, UK, Netherlands and France, $\beta^{(2)}$ for US, UK, Japan and $\beta^{(1)}$ for US, UK and Japan. Note that the estimate of δ takes its biggest value for Japan and its minimum for Italy. Both these values substantially differ from those of the other G7 countries. Comparing the above different sets of SUR based estimates can indicate if imposing homogeneity on the slope and threshold parameters of the model distort the true estimates of the model.

As expected, the results of **Table 6** indicate that the estimates of slope parameters $\beta^{(1)}$ and $\beta^{(2)}$, and threshold parameter δ of model (15) are very close to those of the single equation, when we account for full heterogeneity of its parameters. In contrast, when we allow for full homogeneity, then the estimates of $\beta^{(1)}$ and $\beta^{(2)}$, and δ change considerably. The values of AIC reported in the table indicate that this version of the model, which assumes homogeneity of its all parameters across i, does not constitute the best specification of the data. In terms of both parsimony and fitness of the model into the data, the results of **Table 6** indicate that the version of the model which considers a limited degree of homogeneity of its parameters constitute the best specification of the data. This version leads to more efficient estimates of parameters $\beta^{(1)}$ and $\beta^{(2)}$, by exploiting the cross-section (or cross-equation) dimension of the data. It also provides more representative estimates of them and threshold parameter δ

Summing up, the results of this section clearly show the single equation estimates of the threshold foreign trade multiplier (15) remain robust to the SUR framework. This happens if we allow for possible heterogeneity of both the threshold and slope parameters of the model. Ignoring this heterogeneity can lead to misleading estimates of both the slope and threshold parameters of the threshold model. Both the estimates of slope and threshold parameters seem to be very sensitive to it.

6 Conclusions

We suggested a new method to control for the effects of endogeneity of the threshold variable on the estimates of the slope and threshold parameters of single and system threshold regression models. Our method relies on the copula theory and enables us to relax the assumption that the threshold variable is normally distributed. Under Gaussian (or Student-t) copula functions, the copula theory enables us to deal with the problem of endogeneity of the threshold variable in a simple manner, by adjusting the rhs of a threshold regression model by copula transformations of the threshold variable in each regime of the model which are linear functions. The marginal distributions of the threshold variable in each regime of the model employed in the copula transformations it can be estimated non-parametrically conditionally on the value of threshold parameter.

To evaluate the performance of the method and a search estimation procedure suggested by the paper to consistently estimate the threshold and slope parameters of threshold models, we carried out a Monte Carlo study. This is done for single equation and SUR representations. The results indicate that the suggested method can be successfully implemented to deal with the problem of threshold variable endogeneity effects. This is shown even for cases where the threshold variable follows the Student-t, which can resemble many empirical distributions of economic series observed in practice.

As an empirical illustration of our method, we estimated a threshold model of the foreign trade multiplier, capturing the effects of exports on real output growth. These effects may depend on the level (regime) of the real exchange rate volatility, which plays the role of the threshold variable in the model. To answer the above question, we applied our method to the panel of 'G7' countries using observations from 1963 to 2015. We also suggest a bootstrap testing procedure to formally examine if there are significant differences in the foreign trade multiplier effects across the two regimes of the model, under potential endogeneity of the threshold variable. Our results suggest that, for most countries examined, the foreign trade multiplier depends, crucially, on the high or low volatility regime of the real effective exchange rate. For countries like the US, UK, Japan and France, the foreign trade multiplier exceeds unity when the real exchange rate volatility is at its high-level. This result is also supported by SUR based estimates, allowing for a limited degree of homogeneity in the parameters of the foreign trade threshold model estimated.

References

[1] Aghion, P., P. Bacchetta, R. Ranciere and K. S. Rogoff (2009), Exchange rate volatility and productivity growth: The role of financial development, *Journal of Monetary Economics*, 56, 494–513.

[2] Bagella, M., L. Becchetti and I. Hasan (2006), Real Effective Exchange rate Volatility and Growth: A Framework to Measure Advantages of Flexible vs Costs of Volatility, *Journal of Banking and Finance*, 30, 1149–1169.

[3] Bailey, M. J., G. Tavlas, and M. Ulan (1987), The Impact of Exchange Rate Volatility on Export Growth: Some Theoretical Considerations and Empirical Results, *Journal of Policy Modeling*, 9, 225–243.

[4] Berman, N. and A. Berthou, (2009), Financial Market Imperfections and the Impact of Exchange Rate Movements on Exports, *Review of International Economics*, (17), 103–120

[5] Berman, N and P. Martin, (2012), How do Different Exporters React to Exchange Rate Changes?, *Quarterly Journal of Economics*, 127(1), 437–492.

[6] Caner, M., and B. Hansen (2004), Instrumental Variables Estimation of a Threshold Model, *Econometric Theory*, 20, 813–843.

[7] Chan, K. S., (1993), Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model, *Annals of Statistics*, 21, 520–533.

[8] Darvas, Z., (2012), Real Effective Exchange Rates for 178 Countries: A New Database, Bruegel Working Paper, 2012/06.

[9] Davig, T. and E. M. Leeper (2007), Generalizing the Taylor Principle, *American Economic Review*, 97, 607–635.

[10] Doornik, J. A. and H. Hansen (2008), An Omnibus test for Univariate and Multivariate Normality, *Oxford Bulletin of Economics and Statistics*, 70, 927–939.

[11] Edwards, S. and E. L. Yeyati (2005), Flexible Exchange Rates as Shock Absorbers, *European Economic Review*, 49, 2074–2105.

[12] Egert., B., and A. Morales-Zumaquero (2008), Exchange Rate Regimes, Foreign Exchange Volatility and Export Performance in Central and Eastern Europe: Just Another Blur Project?, *Review of Development Economics*, 12, 577–593.

[13] Fan, Y. and Patton A. J (2014), Copulas in Econometrics, *Annual Review of Economics*, 6(8), 179–200.

[14] Franke, G. (1991), Exchange rate volatility and international trading strategy, *Journal of International Money and Finance*, 10, 292–305.

[15] Frankel, J., and S. Wei (1993), Emerging Currency Blocs. NBER Working Paper No. 4335.

[16] Gonzalo, J. and J. Pitarakis (2002), Estimation and Model Selection Based Inference in Single and Multiple Threshold Models, *Journal of Econometrics*, 110, 319–352.

[17] Gospodinov, N. (2005), Testing for Threshold Nonlinearity in Short-Term Interest Rates, *Journal of Financial Econometrics*, 3, 344–371.

[18] Grier, K. B. and A. D. Smallwood (2013), Exchange Rate Shocks and Trade: A multivariate GARCH-M Approach, *Journal of International Money and Finance*, 37, 282–305.

[19] Griffin, J. M., F. Nardari and R. M. Stultz (2007), Do Investors Trade More When Stocks Have Performed Well? Evidence from 46 Countries, *Review of Financial Studies*, 20, 905–951.

[20] Hansen, B. E, (1996), Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis, *Econometrica*, 62, 413–430.

[21] Hansen, B. E. (2000), Sample Splitting and Threshold Estimation, *Econometrica*, 68, 575–603.

[22] Kapetanios, G. (2001), Model Selection in Threshold Models, *Journal of Time Series Analysis*, 22(6), 733-75.

[23] Kapetanios, G. (2010), Testing Exogeneity in Threshold Models. *Econometric Theory*, 26(1),231-259.

[24] Karlis, D. and Nikoloulopoulos A. K (2009), Modelling multivariate count data using copulas, *Communications in Statistics - Simulation and Computation*, 39, 172–187.

[25] Kazanas, T. and E. Tzavalis (2015), Unveiling the ECB's Monetary Policy Behaviour Under Different Inflation Regimes, *Economica*, 82, 912–937.

[26] Kourtellos, A., T. Stengos and C. M. Tan (2013), The Effect of Public Debt on Growth in Multiple Regimes, *Journal of Macroeconomics*, 2013, 35–43.

[27] Kourtellos, A., T. Stengos and C. M. Tan (2016), Structural Threshold Regression, *Econometric Theory*, 32, 4, 827–860.

[28] Kourtellos, A., T. Stengos and Y. Sun (2017), Endogeneity in Semiparametric Threshold Regression, WP University of Guelph, Canada.

[29] Kraay, A. (2012), How Large is the Government Spending Multiplier? Evidence from World Bank Lending, *Quarterly Journal of Economics*, 127, 829–887.

[30] Kroner, K. and W. Lastrapes (1993), The Impact of Exchange Rate Volatility on International Trade: Reduced form Estimates Using the GARCH-in-Mean Model, *Journal of International Money and Finance*, 12, 298–318.

[31] Lego, B., T. Gebremedhin and B. Cushing (2000), A Multi-Sector Export Base Model of Long-Run Regional Employment, *Growth Agricultural and Resource Economics Review*, 192–197.

[32] Lundbergh, S., T. Teräsvirta, and D. Van Dijk, (2003), Time–Varying Smooth transition Autoregressive Models, *Journal of Business and Economic Statistics*, 21, 104–121.

[33] Meligkotsidou L., E. Tzavalis and I. D Vrontos (2014), A Bayesian method of distinguishing unit root from stationary processes based on panel data models with cross-sectional dependence, *Statistics and Computing*, 24, 297–315.

[34] Patton, A. J (2006), Modelling Asymmetric Exchange Rate Dependence, *International Economic Review*, 47, 527–556.

[35] Patton A. J. (2012), A Review of Copula Models for Economic Time Series, *Journal of Multivariate Analysis*, 110, 4–18.

[36] Pitarakis, J. (2008), Threshold autoregression with a unit root revised, *Econometrica*, 76,1207–1217.

[37] Potter, S. M. (1995), A Nonlinear Approach to US GNP, *Journal of Applied Econometrics*, 10, 109–125.

[38] Reinhart, C. M. and K. S. Rogoff (2010), Growth in a Time of Debt, *American Economic Review*, 100, 573–578.

[39] Robinson, R. (1952), A graphical analysis of the foreign trade multiplier, *Economic Journal*, 62, 546–564

[40] Rothfelder, M. and O. Boldea (2016), Testing for a Threshold in Models with Endogenous Regressors, DP, 2016–029, Tilburg, Center for Economic Research.

[41] Shapiro, S. S. (1980), How to test Normality and other Distributional Assumptions, The ASQC basic references in quality control: *Statistical Techniques*, 3, 1–78.

[42] Tenreyro, S. (2007), On the trade impact of nominal exchange rate volatility, *Journal of Developed Economics*, 82, 485–508.

[43] Teräsvirta, T., Tjøstheim, D. and Granger, C. W. J. (2010), *Modelling Nonlinear Economic Time Series*, Oxford, Oxford University Press.

[44] Tong, H. (2010), Threshold Models in Time Series Analysis - 30 Years on, Working Paper, 471, The University of Hong Kong.

[45] Tsay, R. S., (2010), *Analysis of Financial Time Series*, Second edition, Wiley.

[46] Yeyati, L. E., and F. Sturzenegger (2003), To Float or to Fix: Evidence on the Impact of Exchange Rate Regimes, *American Economic Review*, 93, 1173–1193.

[47] Yu, P. and P. C. B. Phillips (2017), Threshold Regression with Endogeneity, *Journal of Econometrics*, forthcoming.

A Proof of (6)

Using the bivariate copula function (4) and $\varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon_i})$, the conditional pdf $f(\varepsilon_{it}|z^{(h)}_{it}) = c^{(h)}(U_{\varepsilon_{it}}, U_{z^{(h)}_{it}})f(\varepsilon_{it})$, can be written as follows:

$$f(\varepsilon_{it}|z^{(h)}_{it}) = \frac{1}{\sqrt{1-(\rho^{(h)})^2}} \exp\left\{-\frac{(\rho^{(h)})^2(\varepsilon^{*2}_{it} + z^{(*h)2}_{it}) - 2(\rho^{(h)})^2\varepsilon^*_{it}z^{*(h)}_{it}}{2\left(1-(\rho^{(h)})^2\right)}\right\} \frac{1}{\sigma_{\varepsilon_i}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\frac{\varepsilon^2_{it}}{\sigma^2_{\varepsilon_i}}\right\}$$

Equation (**6**) can be derived from the last relationship, by rearranging terms and using the result $\varepsilon_{it} = \Phi^{-1}_{\varepsilon_i}(\Phi(\varepsilon^*_{it})) = \sigma_{\varepsilon_i}\varepsilon^*_{it}$. ■

B Proof of (12)

Using the multivariate copula function (9a) and $\varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon_i})$, the conditional density function $f(\varepsilon_t|z^{(h)}_t) = \frac{f(\varepsilon_t, z_t|z^{(h)}_t)}{f(z^{(h)}_t)}$, can be written as follows:

$$\begin{aligned}
f(\varepsilon_t|z_t \le \delta) &= \frac{f(\varepsilon_t, z_t|z_t \le \delta)}{f(z_t \le \delta)} = c^{(1)}(U_{\varepsilon_{1t}}, ...U_{\varepsilon_{Nt}}, U_{z^{(1)}_{1t}}, ..., U_{z^{(1)}_{Nt}})\prod_{i=1}^{N} f(\varepsilon_{it}) \\
&= \frac{1}{|R^{(h)}|^{1/2}} \exp\left(-\frac{1}{2}U^{*(h)\prime}_t(R^{(h)-1} - I)U^{*(h)}_t\right) \prod_{i=1}^{N} \frac{1}{\sigma_{\varepsilon_i}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\frac{\varepsilon^2_{it}}{\sigma^2_{\varepsilon_i}}\right\} \\
&= \frac{1}{|R^{(h)}|^{1/2}\left(\sqrt{2\pi}\right)^N \prod_{i=1}^{N}\sigma_{\varepsilon_i}} \exp\left\{-\frac{1}{2}U^{*(h)\prime}_t(R^{(h)-1} - I)U^{*(h)}_t\right\} \prod_{i=1}^{N} \exp\left\{-\frac{1}{2}\frac{\varepsilon^2_{it}}{\sigma^2_{\varepsilon_i}}\right\} \\
&= \frac{1}{|R^{(h)}|^{1/2}\left(\sqrt{2\pi}\right)^N |\Sigma_\varepsilon|^{1/2}} \exp\left\{-\frac{1}{2}U^{*(h)\prime}_t(R^{(h)-1} - I)U^{*(h)}_t\right\} \exp\left\{-\frac{1}{2}\varepsilon'_t\Sigma^{-1}_\varepsilon\varepsilon_t\right\}, \\
&= \frac{1}{|R^{(h)}|^{1/2}\left(\sqrt{2\pi}\right)^N |\Sigma_\varepsilon|} \exp\left\{-\frac{1}{2}\left(U^{*(h)\prime}_t R^{(h)-1}U^{*(h)}_t - U^{*(h)\prime}_t U^{*(h)}_t + \varepsilon'_t\Sigma^{-1}_\varepsilon\varepsilon_t\right)\right\}
\end{aligned} \tag{B.1}$$

where $\Sigma_\varepsilon \equiv diag[\sigma^2_{\varepsilon_i}]$. Using $U^{*(h)\prime}_t = (\varepsilon^{*\prime}_t \vdots z^{*(h)\prime}_t)$ and since $\varepsilon_t = \Phi^{-1}_\varepsilon(\Phi(\varepsilon^*_t)) = \Sigma^{1/2}_\varepsilon\varepsilon^*_t$, the quadratic form entering (B.1) can be written as

$$\begin{aligned}
&U^{*(h)\prime}_t R^{(h)-1}U^{*(h)}_t - U^{*(h)\prime}_t U^{*(h)}_t + \varepsilon'_t\Sigma^{-1}_\varepsilon\varepsilon_t \\
&= \varepsilon^{*\prime}_t R_{\varepsilon^*\varepsilon^*}\varepsilon^*_t + \varepsilon^{*\prime}_t R_{\varepsilon^* z^{*(h)}} z^{*(h)}_t + z^{*(h)}_t R_{z^{*(h)}\varepsilon^*}\varepsilon^*_t + z^{*(h)\prime}_t R_{z^{*(h)}z^{*(h)}} z^{*(h)}_t - \varepsilon^{*\prime}_t\varepsilon^*_t - z^{*(h)\prime}_t z^{*(h)}_t + \varepsilon'_t\Sigma^{-1}_\varepsilon\varepsilon_t \\
&= \varepsilon^{*\prime}_t R_{\varepsilon^*\varepsilon^*}\varepsilon^*_t + \varepsilon^{*\prime}_t R_{\varepsilon^* z^{*(h)}} z^{*(h)}_t + z^{*(h)}_t R_{z^{*(h)}\varepsilon^*}\varepsilon^*_t + z^{*(h)\prime}_t R_{z^{*(h)}z^{*(h)}} z^{*(h)}_t - z^{*(h)\prime}_t z^{*(h)}_t
\end{aligned} \tag{B.2}$$

since $\varepsilon_t = \Phi_\varepsilon^{-1}(\Phi(\varepsilon_t^*)) = \Sigma_\varepsilon^{1/2}\varepsilon_t^*$. Using results on quadratic forms for symmetric matrices, like matrix $R^{(h)}$, we can derive the following result:

$$\varepsilon_t^{*\prime} R_{\varepsilon^*\varepsilon^*}\varepsilon_t^* + \varepsilon_t^{*\prime} R_{\varepsilon^* z^{*(h)}} z_t^{*(h)} + z_t^{*(h)} R_{z^{*(h)}\varepsilon^*}\varepsilon_t^* + z_t^{*(h)\prime} R_{z^{*(h)}z^{*(h)}} z_t^{*(h)}$$
$$= z_t^{*(h)\prime} R_{z^{*(h)}z^{*(h)}} z_t^{*(h)} +$$
$$\left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \left(R_{\varepsilon^*\varepsilon^*} - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\right)^{-1} \left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)$$
$$= z_t^{*(h)\prime} R_{z^{*(h)}z^{*(h)}} z_t^{*(h)} +$$
$$\left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \left(R_{\varepsilon^*\varepsilon^*} - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\right)^{-1} \left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)$$
$$= z_t^{*(h)\prime} R_{z^{*(h)}z^{*(h)}} z_t^{*(h)} +$$
$$\left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \left(R_{\varepsilon^*\varepsilon^*} - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\right)^{-1} \left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R^{-1}_{z^{*(h)}z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)$$
$$= z_t^{*(h)\prime} z_t^{*(h)} + \left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \left(R_{\varepsilon^*\varepsilon^*} - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\right)^{-1} \left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)$$
$$= z_t^{*(h)\prime} z_t^{*(h)} + \left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \left(R_{\varepsilon^*\varepsilon^*} - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\right)^{-1} \left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right),$$
$$= z_t^{*(h)\prime} z_t^{*(h)} +$$
$$\left(\varepsilon_t^* - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \Sigma_\varepsilon^{1/2}\Sigma_\varepsilon^{-1/2} \left(R_{\varepsilon^*\varepsilon^*} - R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\right)^{-1} \left(\Sigma_\varepsilon^{1/2}\varepsilon_t^* - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right), \tag{B.3}$$
$$= z_t^{*(h)\prime} z_t^{*(h)} +$$
$$\left(\Sigma_\varepsilon^{1/2}\varepsilon_t^* - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \left(\Sigma_\varepsilon^{1/2} R_{\varepsilon^*\varepsilon^*}\Sigma_\varepsilon^{1/2} - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\Sigma_\varepsilon^{1/2}\right)^{-1} \left(\Sigma_\varepsilon^{1/2}\varepsilon_t^* - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\right.$$
$$= z_t^{*(h)\prime} z_t^{*(h)} +,$$
$$\left(\varepsilon_t - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right)' \left(\Sigma_\varepsilon^{1/2} R_{\varepsilon^*\varepsilon^*}\Sigma_\varepsilon^{1/2} - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*}\Sigma_\varepsilon^{1/2}\right)^{-1} \left(\varepsilon_t - \Sigma_\varepsilon^{1/2} R_{\varepsilon^* z^{*(h)}} R_{z^{*(h)}\varepsilon^*} z_t^{*(h)}\right) \tag{B.4}$$

since $\varepsilon_t = \Phi_\varepsilon^{-1}(\Phi(\varepsilon_t^*)) = \Sigma_\varepsilon^{1/2}\varepsilon_t^*$. Since $R_{z^{*(h)}z^{*(h)}} = I_N$, note that $|R_{z^{*(h)}\varepsilon^*}|$ can be written as follows:

$$|R^{(h)}| = |R_{z^{*(h)}z^{*(h)}}| \left|R_{\varepsilon^*\varepsilon^*} - R_{z^{*(h)}\varepsilon^*} R^{-1}_{z^{*(h)}z^{*(h)}} R_{\varepsilon^* z^{*(h)}}\right| = \left|R_{\varepsilon^*\varepsilon^*} - R_{z^{*(h)}\varepsilon^*} R_{\varepsilon^* z^{*(h)}}\right|$$

Using the last result and noticing $|\Sigma_\varepsilon| = |\Sigma_\varepsilon^{1/2}||\Sigma_\varepsilon^{1/2}|$, since $\Sigma_\varepsilon = diag[\sigma^2_{\varepsilon_i}]$, $|R^{(h)}|^{1/2}|\Sigma_\varepsilon|^{1/2}$ can be written as,

$$\begin{aligned} |R^{(h)}|^{1/2}|\Sigma_\varepsilon|^{1/2} = & \left(|R^{(h)}||\Sigma_\varepsilon|\right)^{1/2} \\ = & \left(|\Sigma_\varepsilon^{1/2}||R^{(h)}||\Sigma_\varepsilon^{1/2}|\right)^{1/2} = \left|\Sigma_\varepsilon^{1/2} R_{\varepsilon^*\varepsilon^*}\Sigma_\varepsilon^{1/2} - \Sigma_\varepsilon^{1/2} R_{z^{*(h)}\varepsilon^*} R_{\varepsilon^* z^{*(h)}}\Sigma_\varepsilon^{1/2}\right|^{1/2} \end{aligned} \tag{B.5}$$

Substituting (B.4)-(B.2) into (B.1) implies relationship (12). ■