Estimation of the Generalized Endogenous Threshold Model

By Hung-Pin Lai and Chang-ching Lin

Department of Economics
National Chung Cheng University

Institute of Economics
Academia Sinica

Oct. 8th, 2012

National Chung Cheng University

Threshold Model (1)

Threshold model:

$$y_i = \begin{cases} \alpha_1 + x_i\beta_1 + e_i, & q_i \leq \gamma, \\ \alpha_2 + x_i\beta_2 + e_i, & q_i > \gamma. \end{cases}$$
$$= (\alpha_1 + x_i\beta_1)\mathbf{1}(q_i \leq \gamma) + (\alpha_2 + x_i\beta_2)\mathbf{1}(q_i > \gamma) + e_i,$$

with sample split rule

$$q_i = \mu + z_i\theta + u_i, \qquad i = 1, \ldots, n.$$

where α_j and β_j, $j = 1, 2$, denote the intercept and slope coefficient of regime j, x_i is the $1 \times k$ vector of exogenous variables, q_i denotes a threshold variable, $e_i \sim iid(0, \sigma_i^2)$ $u_i \sim iid(0, 1)$ and at least one variable in z_i is not in x_i (exclusive condition).

Differences between Sample Selection and Threshold models

1. Limited sample vs full sample
2. Known membership vs Unknown membership (observed threshold variable)

Threshold Model (2)

- Threshold model is widely applied in economics studies due to its nonlinearity and simplicity.
  - Hansen (2000) provides theoretical properties of the threshold estimator for exogenous regressors.
  - Caner and Hansen (2004) consider an IV threshold model for endogenous regressors.
- These threshold models are assumed that the threshold variable is exogenous $\rightarrow$ e_i and q_i are correlated/dependence.

Motivations

Some empirical examples for the ET models can be found in the literature:

- production technology in the manufacture industry: capital-intensive or labor-intensive technology (Lai, 2012)
- hotel services: accommodation and/or restaurant (Lai, 2012)
- parameter heterogeneity by utilizing the threshold model with many fundamental determinants as threshold variables, while institutions might be endogenous (see, for example, Kourtellos, et al. 2010)

There might be some unobserved factors affecting the threshold variable and outcome measure simultaneously and, therefore, they might be endogenous.

Literature Reviews (1)

- Kourtellos, Stengos, and Tan (2011, KTS) treat the problem of an endogenous threshold variable as the problem of omitting selection correction term by imposing the normality and then use IVs with Heckman two-stage regression. Their method:
  1. is simple and easy to interpret
  2. may have biased results when the assumption of normality is invalid or due to tail dependence.
- Lai and Lin (2010) apply the local linear approximation of Fan (1992, 1993) and Ruppert and Wand (1994) to control for the effects of the nonlinear sample correction terms. Their approach
  1. is a semi-parametric method and more flexible on distribution and dependence structure.
  2. may not be as efficient as the parametric estimator since the parametric information is not fully incorporated.

Literature Reviews (2)

- The motivation behind the endogenous threshold (ET) model is similar to that in Heckman's (1979) sample selection (SS) model.
- SS Model with copula approach:
  1. Olsen (1980) and Lee (1982): two-step approach with an assumption of linear relationship between disturbances.
  2. Lee (1983) uses a bivariate normal copula to describe sample selection mechanism when the margins are not normally distributed.
  3. Smith (2003) compares the empirical results from Lee's approach and from the selection models with Archimedean copulas.
  4. Genius and Strazzera (2008); Bhat and Eluru (2009);
  5. Lai, Polachek, and Wang (2009)– stochastic frontier model with sample selection.
  6. Luechinger, Stutzer, and Winkelmann (2010)– switching regression model.

This paper...

We apply a copula approach to estimate the endogenous threshold model under a general specification of dependence.

- The potential contributions of this paper:
  1. The first attempt in the literature to apply a copula approach to handle an endogenous transition variable in the threshold model with non-normally distributed data.
  2. The copula approach utilizes a priori information of the margins and the conjecture of the dependence structure. Thus, it would be more reliable than Kourtellos et al.(2009) and Lin and Lai (2010).
  3. Following the tradition of the Heckman approach, our two-stage nonlinear least squares threshold estimation is easy to implement.
- We provide asymptotic properties of our threshold and slope coefficient estimators to fill a gap in the literature on estimating the threshold model with an endogenous transition variable.

Problem Setup

Consider the following threshold regression model:

$$
\begin{aligned}
y_i &= (\eta_1 + x_i\beta_1)\mathbf{1}(q_i \leq \gamma) + (\eta_2 + x_i\beta_2)\mathbf{1}(q_i > \gamma) + e_i, \\
q_i &= \mu + z_i\theta + u_i,
\end{aligned}
$$

such that at least one variable in z_i is not in x_i (exclusive condition) and e_i and u_i are correlated.

Given the model, the conditional expectation of e_i can be written as

$$\begin{aligned} E\left[e_i|x_i, z_i, q_i\right] &= E\left[e_i|z_i, q_i\right] \\ &= E\left[e_i|z_i, q_i \leq \gamma\right] 1(q_i \leq \gamma) + E\left[e_i|x_i, z_i, q_i > \gamma\right] 1(q_i > \gamma) \\ &= E\left[e_i|u_i \leq \gamma - (\mu + z_i\theta)\right] 1(q_i \leq \gamma) \\ &\qquad + E\left[e_i|u_i > \gamma - (\mu + z_i\theta)\right] 1(q_i > \gamma). \end{aligned}$$

Therefore, the conditional expectation of e_i given the information (x_i, z_i, q_i, γ) can also represented as

$$E\left[e_i|x_i, z_i, q_i, \gamma\right] = \begin{cases} g_1(\gamma - \mu - z_i\theta), & \text{if } q_i \leq \gamma \\ g_2(\gamma - \mu - z_i\theta), & \text{if } q_i > \gamma \end{cases}$$

The threshold model given above can be summarized as

$$\begin{aligned} y_i &= [\alpha_1 + x_i\beta_1 + g_1(\gamma - \mu - z_i\theta)]\,1(q_i \le \gamma) \\ &\qquad + [\alpha_2 + x_i\beta_2 + g_2(\gamma - \mu - z_i\theta)]\,1(q_i > \gamma) + \varepsilon_i, \\ q_i &= \mu + z_i\theta + u_i, \end{aligned}$$

where

$$\varepsilon_i = e_i - E\left[e_i|x_i, z_i, q_i \le \gamma\right]1(q_i \le \gamma) - E\left[e_i|x_i, z_i, q_i > \gamma\right]1(q_i > \gamma).$$

Classic Case

e_i and u_i are jointly normally distributed and $corr(e_i, u_i) = \rho$. Then,

$$\begin{aligned} g_1(q_i, z_i; \mu, \theta, \gamma) &= \frac{-\rho\phi(\gamma - \mu - z_i\theta)}{\Phi(\gamma - \mu - z_i\theta)}, \\ g_2(q_i, z_i; \mu, \theta, \gamma) &= \frac{\rho\phi(\gamma - \mu - z_i\theta)}{1 - \Phi(\gamma - \mu - z_i\theta)}, \end{aligned}$$

which are the inverse Mill's ratios.

Vella (1992, 1998) and Wooldridge (1998), however, have pointed out that deviation from the normality assumption may results in the estimation bias in $(\alpha_j, \beta_j^\top)$, $j = 1, 2$.

Copula Approach

Notice that

- marginal distributions tend to be more tractable than joint distributions.
- some copulas allow for asymmetric tail dependence, which is a particularly important feature in financial data.
- the copula approach is invariant under strictly increasing transformations of the random variables (such as log and exp).
- some statistics methods, such as the tests of Fermanian (2005) and Chen (2007), enable researchers to do model selection and re-specification a posteriori among some candidate copulas.
- it is easy to implement estimation under a copula approach if the data are non-normal.

Copula Representation

Let $F_s(\cdot)$ and $f_s(\cdot)$ be the cdf and pdf of the variable s, where $s = e,\ u$. Let $F_{eu}(e, u)$ denote the true joint cdf of e and u, and δ be the dependent parameter in the corresponding copula function.
According to the Sklar (1959) theorem, there exists a copula function $C(\cdot)$, such that for all $(e_i, u_i) \in \mathrm{R} \times \mathrm{R}$,

$$F_{eu}(e_i, u_i) = C(F_e(e_i), F_u(u_i); \delta),$$

and $C(\cdot)$ is unique if both $F_e(e)$ and $F_u(u)$ are continuous.

The associated joint density is

$$f_{eu}\left(e, u\right) = c_{eu}\left(F_e(e),\ F_u(u)\right) \frac{\partial F_e(e)}{\partial e} \frac{\partial F_u(u)}{\partial u},$$

where

$$c_{eu}\left(F_e(e),\ F_u(u)\right) = \frac{\partial^2 C\left(F_e(e),\ F_u(u)\right)}{\partial F_e(e) \partial F_u(u)}.$$

Given a threshold value γ, the sample will be divided into two groups with $\{i : u_i > \gamma - \mu - z_i\theta\}$ and $\{i : u_i \leq \gamma - \mu - z_i\theta\}$. In the first case, the joint probability of $e < e_i$ and $u_i > \gamma - \mu - z_i\theta$ can be written as

$$\begin{aligned}\Pr(e < e_i, u_i > \gamma - \mu - z_i\theta) &= F_e(e_i) - F_{eu}(e_i, \gamma - \mu - z_i\theta) \\ &= F_e(e_i) - C(F_e(e_i), F_u(\gamma - \mu - z_i\theta); \delta).\end{aligned}$$

It follows that the conditional probability of $e < e_i | u_i > \gamma - \mu - z_i\theta$ is

$$F_{e|u}(e_i | u_i > \gamma - \mu - z_i\theta) = \frac{F_e(e_i) - C(F_e(e_i), F_u(\gamma - \mu - z_i\theta); \delta)}{1 - F_u(\gamma - \mu - z_i\theta)}.$$

Taking the derivative of with respect to e, one gets the conditional pdf of $e_i|u_i > \gamma - \mu - z_i\theta$,

$$f_{e|u}(e_i|u_i > \gamma - \mu - z_i\theta) = \frac{f_e(e_i) - \frac{\partial C(F_e(e_i), F_u(\gamma-\mu-z_i\theta);\delta)}{\partial e}}{1 - F_u(\gamma - \mu - z_i\theta)} = \frac{f_e(e_i)\left[1 - c_e(F_e(e_i), F_u(\gamma - \mu - z_i\theta);\delta)\right]}{1 - F_u(\gamma - \mu - z_i\theta)}.$$

where $c_e = \partial C(F_e(e), F_u(u);\delta)/\partial F_e(e)$.
It suggests that the conditional mean can be written as

$$g_2(z_i;\delta|\mu,\gamma,\theta,C) = \frac{1}{1 - F_u(\gamma - \mu - z_i\theta)}\left[\int_{-\infty}^{\infty} e f_e(e)\left(1 - c_e(F_e(e), F_u(\gamma - \mu - z_i\theta);\delta)\right)\, de\right].$$

Similarly, for the sample with $q_i \leq \gamma$, we have

$$f_{e|u}(e_i|u_i \leq \gamma - \mu - z_i\theta) = \frac{f_e(e_i)\left[c_e(F_e(e_i), F_u(\gamma - \mu - z_i\theta);\delta)\right]}{F_u(\gamma - \mu - z_i\theta)}$$

and

$$g_1(z_i;\delta|\mu,\gamma,\theta,C) = \frac{1}{F_u(\gamma - \mu - z_i\theta)}\left[\int_{-\infty}^{\infty} ef_e(e)\left(c_e(F_e(e), F_u(\gamma - \mu - z_i\theta);\delta)\right)\ de\right].$$

Estimation Procedure

- Step 1, we estimate the regression of q_i by the least squares method and obtain the estimator $(\hat{\mu}, \hat{\theta})$ of (μ, θ) using the data $\{q_i, z_i\}_{i=1}^{n}$.
- Step 2, given a pre-specified copula, we compute $g_1(z_i; \delta|\hat{\mu}, \hat{\theta}, \gamma)$ and $g_2(z_i; \delta|\hat{\mu}, \hat{\theta}, \gamma)$ for all individuals and then estimate $\alpha_1, \alpha_2, \beta_1, \beta_2$ and δ by nonlinear least squares (NLS) estimation for each candidate of γ, denote the estimators $\alpha_1(\gamma), \alpha_2(\gamma), \beta_1(\gamma), \beta_2(\gamma)$ and $\delta(\gamma)$. The objective function in the NLS estimation is defined as

$$\begin{aligned}\mathrm{SSR}(\gamma) &= \sum_{\{i:q_i \le \gamma\}} \left[y_i - \alpha(\gamma) - x_i\beta_1(\gamma) - g_1(z_i; \delta|\gamma, \hat{\mu}, \hat{\theta}, C)\right]^2 \\ &\quad + \sum_{\{i:q_i > \gamma\}} \left[y_i - \alpha(\gamma) - x_i\beta_1(\gamma) - g_2(z_i; \delta|\gamma, \hat{\mu}, \hat{\theta}, C)\right]^2\end{aligned}$$

- Finally, the estimator for γ is defined as the one that minimizes the sum of squared residuals (SSR)

$$\hat{\gamma} = \arg\min_{\gamma \in \Gamma} \mathrm{SSR}(\gamma).$$

Therefore, the NLS estimators for $\alpha_1, \alpha_2, \beta_1, \beta_2$ and δ are the solutions when $\gamma = \hat{\gamma}$.

Some Special Copulas

- Gaussian copula: $C(F_e(e_i), F_u(u_i); \delta) = \Phi_b(\Phi^{-1}(\nu), \Phi^{-1}(\xi), \delta)$, where Φ_b denotes the cdf of the standard bivariate normal distribution, $\nu = \Phi(e)$ and $\xi = \Phi(u)$ and δ is the correlation coefficient in the standard bivariate normal distribution.

The generalized inverse Mill ratios (Lee, 1983) become

$$\begin{aligned} g_1(\gamma - \hat{\mu} - z_i\hat{\theta}; \delta|\Phi_b) &= -\frac{\sigma\delta\varphi(\Phi^{-1}(F_u(\gamma - \hat{\mu} - z_i\hat{\theta})))}{F_u(\gamma - \hat{\mu} - z_i\hat{\theta})} \\ g_2(\gamma - \hat{\mu} - z_i\hat{\theta}; \delta|\Phi_b) &= \frac{\sigma\delta\varphi(\Phi^{-1}(F_u(\gamma - \hat{\mu} - z_i\hat{\theta})))}{1 - F_u(\gamma - \hat{\mu} - z_i\hat{\theta})}. \end{aligned}$$

- The Archimedean copulas:
  1. The Clayton copula: $C(\nu, \xi; \delta) = (\nu^{-\delta} + \xi^{-\delta})^{-1/\delta}$, where $0 < \delta < \infty$ with $\tau = \delta/(\delta + 2)$ and $\varphi(t) = \delta^{-1}(t^{-\delta} - 1)$. The Clayton copula can capture the positive left tail dependence.
  2. The Gumbel copula:
  $C(\nu, \xi; \delta) = \exp\left(-\left[(-\ln(\nu))^{\delta} + (-\ln(\xi))^{\delta}\right]^{1/\delta}\right), 1 \leq \delta < \infty$. The Gumbel copula can capture the positive right tail dependence.
  3. The Frank copula:
  $C(\nu, \xi; \delta) = -\delta^{-1} \ln\left(1 + \frac{(\exp(-\delta\nu)-1)(\exp(-\delta\xi)-1)}{(\exp(-\delta)-1)}\right), -\infty < \delta < \infty$. It is suitable for strong central dependence with very weak tail dependence.

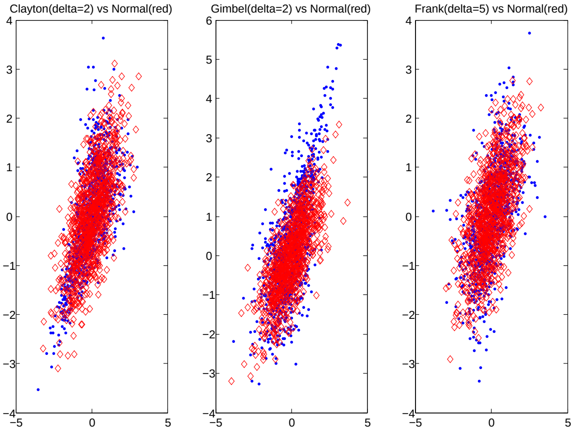

Figure : Plots under Difference Dependence Structures

Assumptions and Theoretical Results

A1 $(x_i, q_i, z_i, e_i, u_i)$ is strictly stationary, ergodic and ρ-mixing, with ρ-mixing coefficients satisfying $\sum_{m=1}^{\infty} \rho_m^{1/2} < \infty$.

A2 $E(e_i|\Im_{i-1}) = 0$ and $E(\varepsilon_i|\gamma - \mu - z_i\theta) = 0$, where $\Im_{i-1}$ is the sigma algebra generated by $(x_{i-j}, z_{i-j}, e_{i-j}, u_{i-j})$.

A3 Let $\chi_i = (x_i, g_{j,i})$, where $g_{j,i} = g_j(z_i; \sigma|\mu, \gamma, \theta, C)$, $j = 1, 2$. Then $E|\chi_i|^4 < \infty$ and $E|\chi_i\varepsilon_i|^4 < \infty$.

A4 For all $\gamma \in \Gamma$, $E\left(|\chi_i|^4 e_i^4 \mid q_i = \gamma\right) \leq M$ and $E\left(|\chi_i|^4 \mid q_i = \gamma\right) \leq M$ for some $M < \infty$, and $f_u(\gamma) \leq \bar{f} < \infty$, where $f_u(\cdot)$ denotes the density function of u_i.

We define the following moments:

$$\kappa(\gamma) = E\left(\chi_i^\top \chi_i d_i(\gamma)\right), \quad \kappa = E\left(\chi_i^\top \chi_i\right),$$

where $d_i(\gamma) = 1(q_i \le \gamma)$. Let χ_{ik} be the k^{th} element of χ_i, and define

$$D(\gamma) = E\left(\chi_i^\top \chi_i | q_i = \gamma\right), \quad D = E\left(\chi_i^\top \chi_i | q_i = \gamma_0\right),$$

and

$$V(\gamma) = E\left(\chi_i^\top \chi_i \varepsilon_i^2 | q_i = \gamma\right), \quad V = E\left(\chi_i^\top \chi_i \varepsilon_i^2 | q_i = \gamma_0\right),$$

where γ_0 represents the true value of γ. Then, we assume:

- A5 $f_u(\gamma)$, $D(\gamma)$, $V(\gamma)$ are continuous at $\gamma = \gamma_0$.
- A6 $\beta_1 - \beta_2 = mn^{-\alpha}$ with $m \neq 0$ and $0 < \alpha < 1/2$.
- A7 $m^\top D(\gamma) m > 0, \quad m^\top V(\gamma) m > 0$ for all $\gamma \in \Gamma$.
- A8 $\bar{\kappa} > |\kappa(\gamma)| > 0$ for all $\gamma \in \Gamma$.
- A9 The copula and marginal distributions are correctively specified.

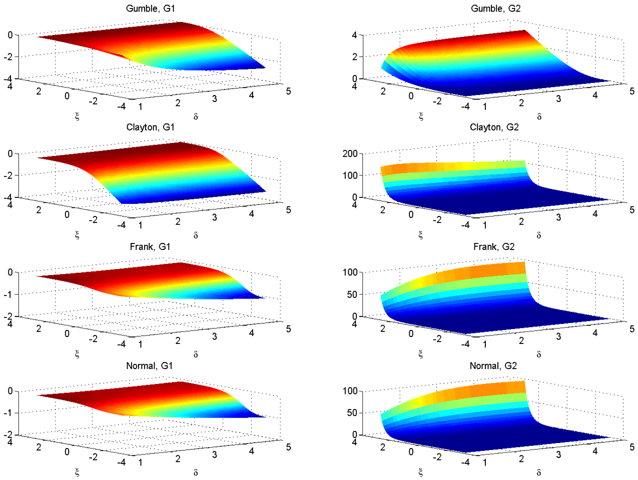

Figure : Monotonicity under Difference Dependence Structures

Asymptotic Properties

Lemma 3.1: *The condition mean function*$g_j(z_i; \sigma|\mu, \gamma, \theta, C)$*is monotonic in* γ.

Theorem 3.2: *Under assumptions [A1]- [A9]*, the *estimator of* γ*is a consistent estimator of* γ_0. *That is* $\hat{\gamma} \xrightarrow{p} \gamma_0$.

Let $\omega = \frac{C^\top VC}{(C^\top DC)^2 f}$, $\quad T = \arg\max_{-\infty<r<\infty} \left[\frac{1}{2}|r| + W(r)\right]$, and $W(r)$ is a two-sided Brownian motion such that

$$W(r) = \begin{cases} W_1(-r)\ , & r < 0, \\ 0\ , & r = 0, \\ W_2(r)\ , & r > 0, \end{cases}$$

where $W_1(-r)$ and $W_2(r)$ are independent standard Brownian motions.

Therefore, based on Theorem 1 of Hansen (2000), we obtain

Theorem 3.3: *Under assumptions [A1]- [A9],* $n^{1-2\alpha}(\widehat{\gamma}_C - \gamma_0) \xrightarrow{d} \omega T$.

Test for the Copula Specification

Let C denote the specified copula function under H_0, and C_0 denote the true copula function. Our objective is to test the following hypothesis:

$$\begin{aligned} H_0 &: \quad F_e(e,u) = C(F_e(e), F_u(u); \delta); \\ H_1 &: \quad F_e(e,u) = C_0(F_e(e), F_u(u); \delta) \end{aligned}$$

The hypothesis can also be reformulated a test for the functional form of $g_j(\cdot)$,

$$\begin{aligned} H_0 &: \quad E\left[e_i | x_i, z_i, q_i\right] = \begin{cases} g_1(\gamma_0 - \mu - z_i\theta; C), & \text{if } q_i \leq \gamma_0; \\ g_2(\gamma_0 - \mu - z_i\theta; C), & \text{if } q_i > \gamma_0. \end{cases} \\ H_1 &: \quad E\left[e_i | x_i, z_i, q_i\right] = \begin{cases} g_1(\gamma_0 - \mu - z_i\theta; C_0), & \text{if } q_i \leq \gamma_0; \\ g_2(\gamma_0 - \mu - z_i\theta; C_0), & \text{if } q_i > \gamma_0. \end{cases} \end{aligned}$$

Hausman Type Test

Let $\hat{\omega}_0$ be the NLS estimator under H_0 and $\hat{\omega}_1$ is a consistent estimator under H_0 and H_1.
By Hausman (1987),

$$Var(\hat{\omega}_1 - \hat{\omega}_0) = Var(\hat{\omega}_1) - Var(\hat{\omega}_0),$$

The Hausman's Chi-squared test statistics has the form

$$W(\gamma) = (\hat{\omega}_1 - \hat{\omega}_0)^{\top} Var(\hat{\omega}_1 - \hat{\omega}_0)^{-1} (\hat{\omega}_1 - \hat{\omega}_0),$$

where k is the dimension of ω.

We follow the work of Hansen (1996) and use the statistic

$$W^* = \sup_{\gamma \in \Gamma} W(\gamma) = \sup_{\gamma \in \Gamma} \left(\hat{\omega}_1 - \hat{\omega}_0\right)^\top Var\left(\hat{\omega}_1 - \hat{\omega}_0\right)^{-1} \left(\hat{\omega}_1 - \hat{\omega}_0\right),$$

which do not require prior knowledge of γ. Since the distribution of W^* is unknown, the bootstrapped distribution of $W^*$is used.

Simulation

DGP: for $i = 1, \ldots, n$,

$$\begin{aligned} y_i &= (\alpha_1 + \beta_1 x_i)1(q_i \le \gamma_0) + (\alpha_2 + \beta_2 x_i)1(q_i > \gamma_0) + e_i, \\ q_i &= \mu + Z_i\theta_1 + u_i, \end{aligned}$$

where $Z = (x_i, z_{1i}, z_{2i})$, x_i, z_{1i}, $z_{2i} \sim iid\ N(0,1)$ and are mutually independent, the true threshold $\gamma_0 = 1$, $(\mu, \theta) = (0, 3, 3)'$, $(\alpha_1, \beta_1, \alpha_2, \beta_2)' = (0.1, 0.2, 0.3, 0.4)'$. To demonstrate the impact of different dependence structures on estimation, we allows the errors e_i and u_i are generate from normal distribution with different dependence structures. For each case in each type, we consider $n = \{50, 100, 200\}$. The number of Monte Carlo replications is 1,000 in each of the experiments.

We consider generate errors with Normal-type, Gumbel-type, Clayton-type and Frank-type dependence structures. (see Marshall and Olkin, 1988; Cherubini, Luciano, and Vecchiato, 2004)

- Normal type dependence errors: are simply generated by letting $u_i \sim iid(0,1)$, $\rho = 0.5$, $e_i = (\rho u_i + (1-\rho)\zeta_i)/\sqrt{\rho^2 + (1-\rho)^2}$, and $\zeta_i \sim iid(0,1)$ and independent of u_i.
- Clayton-type dependence errors are generated by: (2a) generating two mutually independent variables $(\zeta_1, \zeta_2)'$ from $U(0,1)$. (2b) let $\varepsilon_1 = \zeta_1$ and let

$$\varepsilon_2 = \left[\zeta_1^{-\delta}\left(\zeta_2^{\frac{-\delta}{1+\delta}} - 1\right) + 1\right]^{-\frac{1}{\delta}}.$$

(2c) obtain $(u, e)' = (\Phi^{-1}(\varepsilon_1), \Phi^{-1}(\varepsilon_2))'$, where Φ^{-1} denotes the inverse function of a standard normal cumulative distribution function.

- Frank-type dependence errors are generated by: (3a) generating two mutually independent variables $(\zeta_1, \zeta_2)'$ from $U(0,1)$. (3b) let $\varepsilon_1 = \zeta_1$ and let

$$\varepsilon_2 = -\frac{1}{\delta} \ln \left[1 + \frac{\zeta_2 \left(1 - \exp(-\delta)\right)}{\zeta_2 \left(\exp(-\delta\zeta_1) - 1\right) - \exp(-\delta\zeta_1)} \right].$$

  (3c) obtain $(u, e)' = (\Phi^{-1}(\varepsilon_1), \Phi^{-1}(\varepsilon_2))'$.

- Gumbel-type dependence errors are generated by: (4a) generating two mutually independent variables $(\zeta_1, \zeta_2)'$ from $U(0,1)$. (4b) let $\varepsilon_1 = \zeta_1$ and solve $0 \leq \varepsilon_2 \leq 1$ satisfying

$$\zeta_2 = \frac{\varphi(c_2)}{\varphi(c_1)}$$

  with $\varphi(t) = -\delta^{-1} \exp(-t^{1/\delta}) t^{\frac{1-\delta}{\delta}}$, $c_1 = (-\ln(\zeta_1))^{-\delta}$, and $c_2 = c_1 + (-\ln(\zeta_2))^{-\delta}$. (4c) obtain $(u, e)' = (\Phi^{-1}(\varepsilon_1), \Phi^{-1}(\varepsilon_2))'$. We let $\delta = 0.5$ for all different copula dependence structures.

Table : MSEs for normal errors with different dependent structure

	N=50											
	γ	α_1	β_1	α_2	β_2	Coff	γ	α_1	β_1	α_2	β_2	Coff
Copula	GDP: Normal Dep						GDP: Clayton Dep					
OLS	0.502	0.385	0.085	0.425	0.102	0.996	0.509	0.945	0.089	0.274	0.037	1.345
KST	0.458	1.115	0.084	0.638	0.093	1.929	0.430	1.026	0.068	0.290	0.035	1.419
Clayton	0.436	0.170	0.081	0.194	0.102	0.547	0.437	0.242	0.068	0.095	0.037	0.441
Gumbel	0.436	0.168	0.077	0.179	0.099	0.524	0.438	0.255	0.069	0.100	0.036	0.460
Frank	0.470	0.249	0.084	0.309	0.103	0.744	0.483	0.520	0.074	0.171	0.037	0.801
Copula	GDP: Gumbel Dep						GDP: Frank Dep					
OLS	0.837	0.352	0.038	0.848	0.061	1.298	0.572	0.270	0.160	0.248	0.128	0.806
KST	0.744	31.182	0.040	87.373	0.063	118.657	0.550	0.932	0.158	1.413	0.134	2.637
Clayton	0.681	0.153	0.037	0.389	0.069	0.648	0.511	0.200	0.145	0.171	0.112	0.628
Gumbel	0.732	0.133	0.037	0.314	0.060	0.543	0.516	0.211	0.146	0.174	0.114	0.645
Frank	0.774	0.221	0.039	0.593	0.072	0.924	0.553	0.284	0.157	0.262	0.144	0.847

Table : MSEs for normal errors with different dependent structure

	N=100											
	γ	α_1	β_1	α_2	β_2	Coff	γ	α_1	β_1	α_2	β_2	Coff
Copula	GDP: Normal Dep						GDP: Clayton Dep					
OLS	0.349	0.273	0.031	0.287	0.031	0.621	0.380	0.793	0.031	0.202	0.012	1.038
KST	0.312	0.201	0.028	0.187	0.028	0.444	0.354	0.425	0.027	0.085	0.012	0.549
Clayton	0.296	0.108	0.028	0.115	0.028	0.278	0.347	0.149	0.026	0.043	0.012	0.230
Gumbel	0.304	0.102	0.027	0.106	0.027	0.262	0.350	0.163	0.027	0.047	0.012	0.249
Frank	0.327	0.135	0.028	0.143	0.029	0.335	0.368	0.358	0.028	0.091	0.012	0.489
Copula	GDP: Gumbel Dep						GDP: Frank Dep					
OLS	0.730	0.259	0.013	0.756	0.026	1.055	0.426	0.178	0.038	0.153	0.042	0.411
KST	0.646	0.193	0.013	0.395	0.024	0.626	0.411	0.242	0.037	0.216	0.044	0.539
Clayton	0.567	0.101	0.014	0.296	0.024	0.435	0.393	0.126	0.036	0.113	0.041	0.316
Gumbel	0.634	0.084	0.013	0.237	0.025	0.359	0.390	0.121	0.036	0.114	0.041	0.311
Frank	0.679	0.134	0.013	0.440	0.026	0.613	0.422	0.136	0.038	0.140	0.044	0.357

Table : MSEs for normal errors with different dependent structure

	N=200											
	γ	α_1	β_1	α_2	β_2	Coff	γ	α_1	β_1	α_2	β_2	Coff
Errors	GDP: Normal Dep						GDP: Clayton Dep					
OLS	0.225	0.213	0.012	0.219	0.012	0.457	0.334	0.720	0.014	0.178	0.006	0.918
KST	0.201	0.070	0.011	0.073	0.011	0.166	0.300	0.109	0.012	0.033	0.006	0.160
Clayton	0.200	0.066	0.012	0.081	0.012	0.170	0.299	0.075	0.012	0.024	0.006	0.116
Gumbel	0.203	0.059	0.012	0.065	0.011	0.147	0.305	0.079	0.012	0.026	0.006	0.123
Frank	0.205	0.069	0.012	0.083	0.012	0.175	0.328	0.284	0.012	0.069	0.006	0.371
Errors	GDP: Gumbel Dep						GDP: Frank Dep					
OLS	0.636	0.220	0.007	0.663	0.011	0.901	0.264	0.121	0.015	0.115	0.014	0.265
KST	0.552	0.055	0.007	0.170	0.010	0.242	0.237	0.069	0.014	0.060	0.014	0.158
Clayton	0.451	0.071	0.007	0.240	0.011	0.328	0.228	0.065	0.015	0.065	0.014	0.159
Gumbel	0.546	0.048	0.007	0.162	0.010	0.227	0.236	0.057	0.014	0.054	0.014	0.139
Frank	0.550	0.100	0.007	0.342	0.011	0.460	0.238	0.047	0.014	0.063	0.015	0.139

Main findings:

- Failure to control for endogeneity will lead to larger MSE.
- The KST, which has a normal copula might results in very misleading estimates when dependence are not normal.
- When $n = 50$, the methods with Clayton and Gumbel are quite robust to all GDPs and can outperform the KST even when dependence are normal type.
- When n is large (n=200), the method with correct copula can deliver the smallest MSE.

- For robustness, we then demonstrate the impact of non-normally distributed errors on estimation by generating (u_i, e_i) from mixed normal distributions.
- Generate (u_i, v_i) from a (standardized) mixed normal distribution, in which errors are first generated from $N(0,1)$ with probability p and from $N(\mu,1)$ with probability $1-p$, and then are standardized by their sample means and standard errors. u_i and v_i are mutually independent and $e_i = (\rho u_i + (1-\rho)v_i)/\sqrt{\rho^2 + (1-\rho)^2}$ with $\rho = 0.7$.
- In this set of simulations, we particularly interest the accuracy of slope parameters.

Table : MSEs for mixed-normal errors

	γ	β_1	β_2	slope	γ	β_1	β_2	slope
	$(\mu, p) = (4.0.7)$				$(\mu, p) = (4.0.8)$			
OLS	9.053	0.092	0.035	0.127	9.906	0.094	0.036	0.129
KST	4.560	0.030	0.022	0.053	5.696	0.030	0.023	0.053
Clayton	4.315	0.026	0.023	0.049	5.784	0.026	0.024	0.050
Gumbel	4.419	0.029	0.022	0.051	5.503	0.028	0.023	0.050
Frank	5.061	0.039	0.024	0.063	6.514	0.040	0.026	0.066
	$(\mu, p) = (5.0.8)$				$(\mu, p) = (8.0.8)$			
OLS	10.243	0.094	0.036	0.131	10.258	0.096	0.037	0.133
KST	5.626	0.030	0.023	0.054	5.826	0.032	0.024	0.055
Clayton	5.507	0.025	0.025	0.050	5.598	0.026	0.025	0.051
Gumbel	5.574	0.028	0.023	0.051	5.702	0.030	0.023	0.053
Frank	6.345	0.039	0.026	0.065	6.179	0.039	0.027	0.066

Note: “slope” denotes the sum of MSEs over slope parameters (β_1, β_2).

Conclusion

- We propose a two-stage nonlinear least squares method to estimate the endogenous threshold model.
- With the priori information of the margins and proper chosen of the copula, our estimator may avoid the estimation bias resulted from wrongly imposing the assumption on the dependent structure.
- Our simulation results show that our approach can deliver more accurate threshold and slope estimates when the errors have non-normal dependent structures.
- Future Studies:
  1. The maximum likelihood method by the one step estimation. However, it remains to establish the asymptotic properties.
  2. The tests of Fermanian (2005) and Chen (2007) enable researchers to do model selection and re-specification a posteriori among some candidate copulas.

Thank you!