Economics Letters 215 (2022) 110498

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Analytical bias correction for two-step fixed effects models with copula-distributed errors

Costanza Naguib *

University of Bern, WVI, Bern, Schanzeneckstrasse 1, Postfach, 3001 Bern, Switzerland
Università della Svizzera italiana (USI), Faculty of Economics, Lugano, Switzerland

ARTICLE INFO

Article history:
Received 16 August 2021
Received in revised form 31 March 2022
Accepted 4 April 2022
Available online 9 April 2022

JEL classification:
C23

Keywords:
Two-step fixed effects
Analytical bias correction
Relative mobility

ABSTRACT

We derive an analytical bias correction for two-step fixed effects models with copula-distributed errors. We work out the approximate bias correction for the Gaussian copula and present a numerical computation exercise for three other copula families. The results are derived for both n and T large.

1. Introduction

[1] The aim of this paper is to develop an analytical expression for the bias correction for two-step fixed effects models with copula-distributed errors. These models can be used to describe relative mobility, and in particular for studying earnings dynamics (see e.g. Bonhomme and Robin, 2009). There are currently only a few studies which develop an analytical bias correction for two-step models (e.g. Fernandez-Val and Vella, 2011, Cattaneo et al., 2019). We extend their results by developing the expressions for a two-step (parametric) copula model under the asymptotics when both n and T are large.

2. A model for relative mobility

Let us consider the following linear panel model (see e.g. Cameron and Trivedi, 2005):

$$y_{it} = \alpha + x_{it}'\beta + \eta_i + \lambda_t + \varepsilon_{it} \tag{2.1}$$

where y_{it} is the outcome variable (e.g. wage), x_{it} is a vector of exogenous individual explanatory variables, η_i is the individual fixed effect and λ_t is the time fixed effect. We make the following normalizations to ensure that the model is identified: $\sum_{t=1}^{T}\lambda_t = 0$ and $\sum_{i=1}^{n}\eta_i = 0$. For the rest of the paper, we rely on the following Assumptions:

Assumption A.1. Error process (ε_{it}) is stationary and weakly serially dependent over the time dimension t.

Assumption A.2. Errors ε_{it} are independent and identically distributed across the index i.

Assumption A.3. Explanatory variables and error term are not correlated.

Assumption A.3 is needed for consistency of $\hat{\beta}$ in (2.1). The parameter of interest is:

$$\theta_0 = \arg\max_{\theta\in\Theta} E[\log c(U_{it}, U_{i,t-1}; \theta)] \tag{2.2}$$

where $\Theta \subset \mathbb{R}^p$ and $U_{it} = F(\varepsilon_{it})$, $U_{i,t-1} = F(\varepsilon_{i,t-1})$ are the present and past uniform ranks, which have a dynamics given by the copula $c(\cdot;\cdot;\theta)$. $F(\cdot)$ stands for the cumulative distribution function (cdf). We are interested in the following two-step estimator:

$$\hat{\theta} = \arg\max_{\theta\in\Theta} \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T} \log c(\hat{U}_{it}, \hat{U}_{i,t-1}; \theta) \tag{2.3}$$

where $\hat{U}_{it} = \hat{F}(\hat{\varepsilon}_{it})$, $\hat{U}_{i,t-1} = \hat{F}(\hat{\varepsilon}_{i,t-1})$, $\hat{\varepsilon}_{it}$ is the OLS residual estimated in model (2.1) and $\hat{F}$ is its empirical distribution. We consider an asymptotics such that both n and T are large. Under regularity conditions we have $\hat{\theta} \xrightarrow{p} \theta_T$ where:

$$\theta_T = \arg\max_{\theta\in\Theta} \bar{L}_T(\theta), \tag{2.4}$$

* Correspondence to: University of Bern, WVI, Bern, Schanzeneckstrasse 1, Postfach, 3001 Bern, Switzerland.
E-mail address: costanza.naguib@vwi.unibe.ch.

1 I would like to thank Patrick Gagliardini, Blaise Melly, Nicola Gnecco and Kaspar Würtlich for their insightful advice.

https://doi.org/10.1016/j.econlet.2022.110498

$$\bar{L}_T(\theta) := \operatorname*{plim}_{n\to\infty} \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T} \log c(\hat{U}_{it}, \hat{U}_{i,t-1}; \theta).$$

Our goal is to derive an expansion for θ_T when T is large. For generic copula arguments $u, v \in [0, 1]$, where $u = U_{it}$ and $v = U_{i,t-1}$, by applying Taylor approximations we obtain (proof reported in Appendix A):

$$\bar{L}_T(\theta) = \underbrace{E[\log c(u, v; \theta)]}_{\equiv L_0(\theta)} + \frac{1}{T}L_1(\theta) + o\left(\frac{1}{T}\right) \tag{2.5}$$

where:

$$\begin{aligned} L_1(\theta) = {} & -2\sum_{j=-\infty}^{\infty} E\left[\frac{\partial \log c(u, v; \theta)}{\partial u} f(\varepsilon_{it})\varepsilon_{i,t-j}\right] \\ & + E\left[\frac{\partial \log c(u, v; \theta)}{\partial u} f'(\varepsilon_{it}) + \frac{\partial^2 \log c(u, v; \theta)}{\partial u^2}(f(\varepsilon_{it}))^2 \right. \\ & \left. + \frac{\partial^2 \log c(u, v; \theta)}{\partial u \partial v} f(\varepsilon_{it}) f(\varepsilon_{i,t-1})\right] \times \omega^2 \\ & + 2E\left[\frac{\partial \log c(u, v; \theta)}{\partial u} H(\varepsilon_{it})\right]. \end{aligned} \tag{2.6}$$

where $f(\cdot)$ is the probability density function (pdf) and $H(\varepsilon) = \sum_{j=-\infty}^{\infty} E[\varepsilon_{i,t+j}|\varepsilon_{i,t} = \varepsilon] f(\varepsilon) + \frac{\omega^2}{2} f'(\varepsilon)$, and where $\omega^2 = \sum_{-\infty}^{\infty} Cov(\varepsilon_{it}, \varepsilon_{i,t-j})$. This leads us to the following Proposition:

Proposition 1. *The vector θ_T defined in (2.4) is such that for large T:*

$$\theta_T = \theta_0 + \frac{1}{T}B + o\left(\frac{1}{T}\right) \tag{2.7}$$

where θ_0 is the true parameter value and the bias term B is equal to:

$$B = I_0^{-1}\frac{\partial L_1(\theta_0)}{\partial \theta} \tag{2.8}$$

where

$$I_0 = -\frac{\partial^2 L_0(\theta_0)}{\partial\theta\partial\theta'} = -E\left[\frac{\partial^2 \log c(u, v; \theta_0)}{\partial\theta\partial\theta'}\right] \tag{2.9}$$

The derivation of an analytical expression of the bias term has some analogies with Hahn and Newey (2004), pp. 1303–1305), but the maximization criterion is different. In our case, indeed, the individual fixed effects are estimated separately via the linear model in (2.1) and hence do not depend on the parameter vector θ. In the following, we present the explicit formula for the approximate bias correction for the case of a Gaussian copula.

2.1. Approximate bias computation for the Gaussian copula

In this part we compute all the terms that appear in Eq. (2.8) for the case of a parametric Gaussian copula with Gaussian marginals. The Gaussian copula is such that,

$$\log c(u, v; \rho) = -\frac{1}{2}\log(1-\rho^2) + \frac{1}{2}(x^2+y^2) + \frac{2\rho xy - x^2 - y^2}{2(1-\rho^2)},$$

with $x = \Phi^{-1}(u)$, $y = \Phi^{-1}(v)$, $\Phi^{-1}(\cdot)$ is the standard normal quantile function, and $\Phi(\cdot)$ is the standard normal cdf. We make the following assumption.

Assumption A.4. Variables ε_{it} and ε_{it-1} are jointly normal distributed, have both zero mean and unitary variance and have correlation coefficient equal to ρ.

Assumption A.4 implies that:

$$E(\varepsilon_{is}|\varepsilon_{it} = \varepsilon) = \rho^{|s-t|}\varepsilon \tag{2.10}$$

hence,

$$H(\varepsilon) = \omega^2\varepsilon f(\varepsilon) + \frac{\omega^2}{2}f'(\varepsilon) + o\left(\frac{1}{T}\right) = -\frac{\omega^2}{2}f'(\varepsilon) \tag{2.11}$$

By substituting this expression into Eq. (2.6) and simplifying equal terms we obtain:

$$\begin{aligned} L_1(\rho) = {} & -2\sum_{j=-\infty}^{\infty} E\left[\frac{\partial \log c(u, v; \rho)}{\partial u} f(\varepsilon_{it})\varepsilon_{i,t-j}\right] \\ & + E\left[\frac{\partial^2 \log c(u, v; \rho)}{\partial u^2}(f(\varepsilon_{it}))^2 \right. \\ & \left. + \frac{\partial^2 \log c(u, v; \rho)}{\partial u \partial v} f(\varepsilon_{it}) f(\varepsilon_{i,t-1})\right] \times \omega^2. \end{aligned}$$

The partial derivative wrt ρ is:

$$\begin{aligned} \frac{\partial L_1(\rho_0)}{\partial \rho} = {} & -2\sum_{j=-\infty}^{\infty} E\left[\frac{\partial^2 \log c(u, v; \rho_0)}{\partial \rho \partial u} f(\varepsilon_{it})\varepsilon_{i,t-j}\right] \\ & + E\left[\frac{\partial^3 \log c(u, v; \rho_0)}{\partial \rho \partial u^2}(f(\varepsilon_{it}))^2 \right. \\ & \left. + \frac{\partial^3 \log c(u, v; \rho_0)}{\partial \rho \partial u \partial v} f(\varepsilon_{it}) f(\varepsilon_{i,t-1})\right] \times \omega^2. \end{aligned} \tag{2.12}$$

The partial derivatives of the log Gaussian copula density are as follows:

$$\begin{aligned} \frac{\partial \log c(u, v; \rho)}{\partial \rho} &= \frac{\rho}{(1-\rho^2)} + \frac{1}{(1-\rho^2)^2}[xy(1+\rho^2) - \rho(x^2+y^2)] \\ \frac{\partial^2 \log c(u, v; \rho)}{\partial \rho^2} &= \frac{1+\rho^2}{(1-\rho^2)^2} \\ &\quad + \frac{1}{(1-\rho^2)^3}[2\rho(3+\rho^2)xy - (1+3\rho^2)(x^2+y^2)] \\ \frac{\partial^2 \log c(u, v; \rho)}{\partial \rho \partial u} &= \frac{1}{(1-\rho^2)^2\phi(x)}[y(1+\rho^2) - 2\rho x] \\ \frac{\partial^3 \log c(u, v; \rho)}{\partial \rho \partial u^2} &= \frac{1}{(1-\rho^2)^2(\phi(x))^2}[xy(1+\rho^2) - 2\rho(1+x^2)] \\ \frac{\partial^3 \log c(u, v; \rho)}{\partial \rho \partial u \partial v} &= \frac{1}{(1-\rho^2)^2\phi(x)\phi(y)}(1+\rho^2) \\ I_0 &= E\left[-\frac{\partial^2 \log c(u, v; \rho)}{\rho^2}\right] = \frac{1+\rho^2}{(1-\rho^2)^2} \end{aligned}$$

By inserting the expressions computed above for the partial derivatives into Eq. (2.12) we obtain:

$$\frac{\partial L_1(\rho_0)}{\partial \rho} = -\frac{(1+\rho_0^2)(1-\rho_0)\omega^2}{(1-\rho_0^2)^2} \tag{2.13}$$

where $\omega^2 = \frac{(1+\rho_0)}{(1-\rho_0)}$. Hence, the analytical expression for the bias is equal to

$$B = -(1+\rho_0) \tag{2.14}$$

The bias correction would be equal to $-\frac{1}{T}$ in case u and v were independent (i.e. $\rho_0 = 0$). The absolute value of the bias is increasing in the value of ρ_0.

2.2. Numerical computation exercise for other copula families

For copula families different from the Gaussian one, computing the expectations in Eqs. (2.8) and (2.9) is not straightforward. For this reason, in this subsection we present numerical computation results for a few popular copulas, namely the Fairlie–Gumbel–Morgenstern (FMG, Genest and Favre (2007)), the Frank

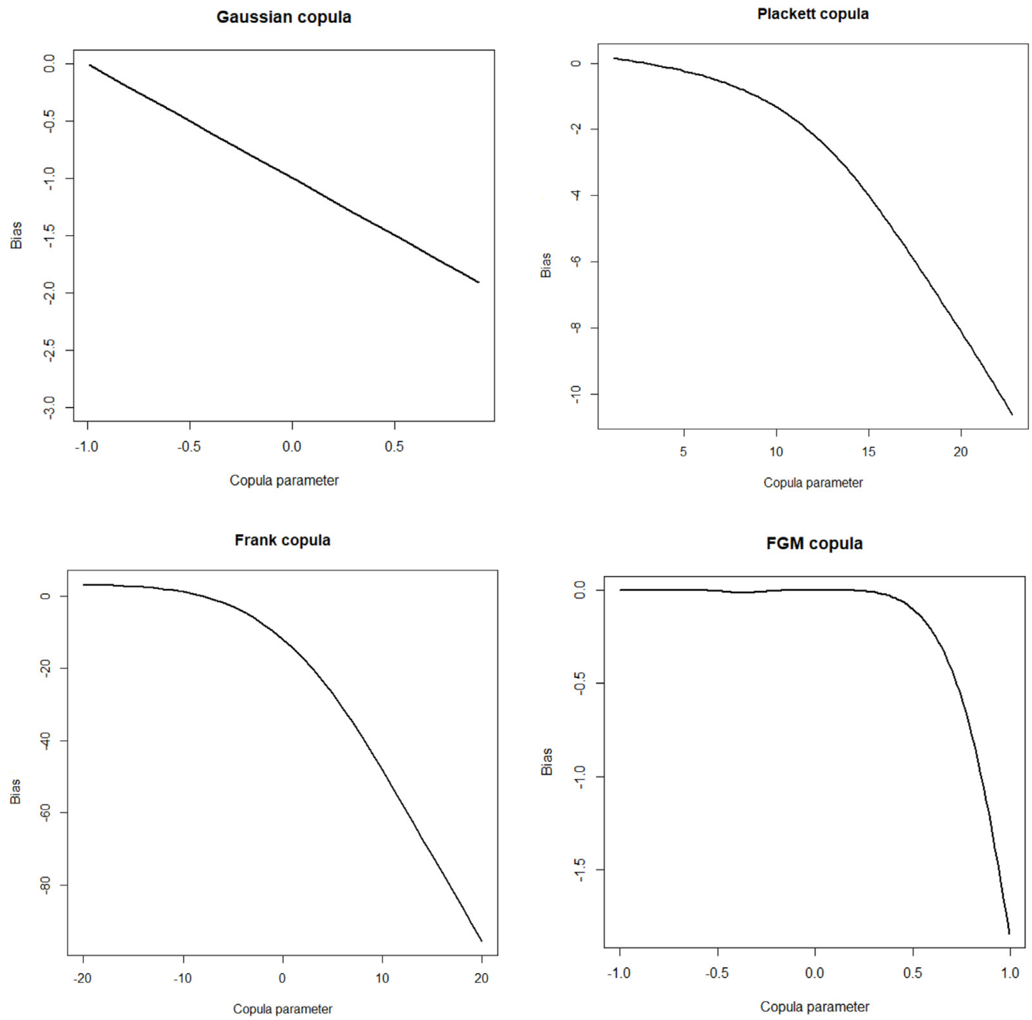

Fig. 1. Approximate bias B for different copula families, as a function of the copula parameter. Upper left panel: Gaussian copula, upper right panel: Plackett copula, bottom left panel: Frank copula, bottom right panel: FGM copula.

copula and the Plackett copula (Joe, 1997; Nelsen, 2007). Throughout this subsection, we assume that variables ε_{it} and $\varepsilon_{i,t-1}$ have both zero mean and unitary variance, and their marginal distribution is Gaussian. For each of the copula families under scrutiny, we simulate 500 observations of the arguments u, v. It is now easy to replace all the theoretical expectations in Eqs. (2.8)–(2.9) with their empirical counterparts (i.e. sample means, sample variances and sample covariances). In each panel of Fig. 1 we report the value of quantity (2.8), numerically computed as explained above, for different copula families. For the Gaussian copula, we simply apply Eq. (2.14).

For the Plackett copula, the copula parameter ranges from 0 to ∞. If this parameter is lower than 1, we witness a negative correlation between the present and past rank, whereas if the parameter is greater than 1, the correlation is positive. The limiting case with the parameter of the Plackett copula equal to 1 corresponds to the case of independence between past and present ranks (Joe, 1997). For the Frank copula, the parameter is $\theta \in \{-\infty; \infty\}\backslash 0$. There is negative dependence for negative values of the parameter and vice versa. The case of independence would correspond to $\theta = 0$, which is however not admissible, as the parameter is at the denominator in the Frank copula cdf.[2]

For the FGM copula, the parameter is: $\theta \in [-1; 1]$, where values smaller than zero stand for negative dependence, values larger than zero stand for positive dependence and $\theta = 0$ stands for independence (Joe, 1997).

From Fig. 1, we deduce that the bias B has a somehow similar pattern for all the copula families considered, i.e. for most parameter values it is increasing in absolute value in the parameter value. However, whereas for the Gaussian and for the FGM copula the bias is always either zero or negative, for the Plackett and for the Frank copula a distinction is needed. In these latest two cases, indeed, when the copula parameter is small enough (i.e. smaller than 2 for the Plackett copula and smaller than around -7 for the Frank copula), the bias is positive and decreasing in the parameter value.

Note that in general the bias B is not zero in the case of independence between u and v. This is because the bias comes from the estimation of the first-step residual in the first stage. In a simple two-period model, the fixed effects for an individual is negative in period $t = 1$ and positive in period $t = 2$, or vice versa, but always summing up to zero for each individual over the two periods. In such a case, our model would detect a negative association between u and v even if there were independence in the second-stage copula. This would entail a negative bias in the estimation of the copula parameter, consistently with the results of our numerical computations reported in Fig. 1.

[2] For the graph, a value slightly above and a value slightly below zero have been used.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econlet.2022.110498.

References

Bonhomme, S., Robin, J.M., 2009. Assessing the equalizing force of mobility using short panels: France, 1990–2000. Rev. Econ. Stud. 76 (1), 63–92.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.

Cattaneo, M.D., Jansson, M., Ma, X., 2019. Two-step estimation and inference with possibly many included covariates. Rev. Econom. Stud. 86 (3), 1095–1122.

Fernandez-Val, I., Vella, F., 2011. Bias corrections for two-step fixed effects panel data estimators. J. Econometrics 163 (2), 144–162.

Genest, C., Favre, A.C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. 12 (4), 347–368.

Hahn, J., Newey, W., 2004. Jackknife and analytical bias reduction for nonlinear panel models. Econometrica 72 (4), 1295–1319.

Joe, H., 1997. Multivariate Models and Multivariate Dependence Concepts. CRC Press.

Nelsen, R.B., 2007. An Introduction to Copulas. Springer Science & Business Media.