Political Analysis (2024), **32**, 417–430
doi:10.1017/pan.2023.45

ARTICLE

Normal Mode Copulas for Nonmonotonic Dependence

Kentaro Fukumoto

Department of Political Science, Gakushuin University, Tokyo, Japan.

Email: Kentaro.Fukumoto@gakushuin.ac.jp

(Received 15 November 2022; revised 4 August 2023; accepted 29 September 2023; published online 13 February 2024)

Abstract
Copulas are helpful in studying joint distributions of two variables, in particular, when confounders are unobserved. However, most conventional copulas cannot model joint distributions where one variable does not increase or decrease in the other in a monotonic manner. For instance, suppose that two variables are linearly positively correlated for one type of unit and negatively for another type of unit. If the type is unobserved, we can observe only a mixture of both types. Seemingly, one variable tends to take either a high or low value (or a middle value) when the other variable is small (large), or vice versa. To address this issue, I consider an overlooked copula with trigonometric functions (Chesneau [2021, *Applied Mathematics*, 1(1), pp. 3–17]) that I name the "normal mode copula." I apply the copula to a dataset about government formation and duration to demonstrate that the normal mode copula has better performance than other conventional copulas.

Keywords: asymmetricity; coalition government; duration; government formation; joint distribution; sine function; trigonometric function

Edited by: Jeff Gill

1. Introduction

When two variables are simultaneously generated, it becomes challenging and promising to study the relationship between them. For instance, if unobserved confounders affect them, regressing one variable on the other leads to endogeneity bias. Scholars may have reasonable substantive knowledge or theory about the marginal distributions of the two variables but no idea about the conditional distribution of one variable given the other. One approach is to directly analyze the joint distribution. When one variable includes some of the same information as the other, analyzing them together will lead to less biased and more efficient estimation, as well as better prediction without assuming selection on observables. Classic methods include the seemingly unrelated regression and Heckman's sample selection models, although they assume only a bivariate normal distribution for the variables.[1] Instead, as I elaborate on below, copula functions are flexible and helpful because they model how dependent the two variables are on each other, whatever marginal distribution each variable follows. In political science, Braumoeller *et al.* (2018), Chiba, Martin, and Stevenson (2015), Chiba, Metternich, and Ward (2015), and Fukumoto (2015) have employed copulas.[2] In finance, Li (2000) improves credit derivative valuation by accounting for the default correlation with the help of copulas.

[1] Furthermore, analysts have to integrate a latent variable out of the selection model.
[2] Sartori (2003) modifies Heckman's sample selection model by using the Fréchet upper bound copula in effect, although she does not mention a copula. Gomes *et al.* (2019) explicitly incorporate copulas in their sample selection model.

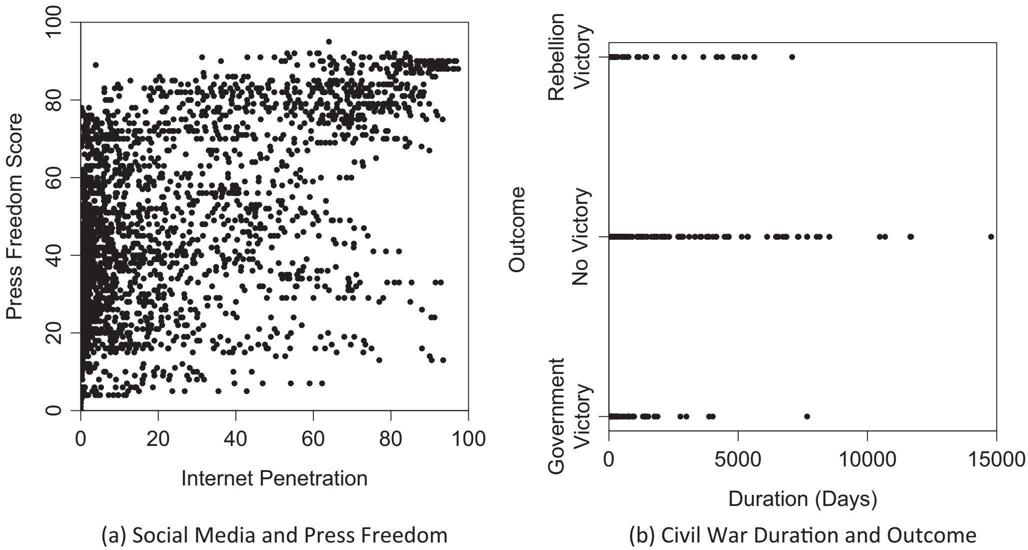

Figure 1. Empirical examples of nonmonotonic dependence. (a) Each dot corresponds to a country in a year, 2000 to 2015 ($n = 2,560$). (b) Each dot corresponds a civil war, 1946 to 2003 ($n = 267$).

However, an understudied shortcoming of copulas is that most conventional copulas cannot model joint distributions where one variable does not increase or decrease in the other in a monotonic manner. To date, little attention has been devoted to such situations because those scenarios "do *not seem* to arise often in applications" (Hofert *et al.* 2018, 173, emphasis added). Nevertheless, they *do* sometimes arise, even in political science. For instance, suppose that two variables are linearly positively correlated for one type of unit and negatively correlated for another type of unit. If the type is unobserved, we can observe only a mixture of both types. Seemingly, one variable tends to take either a high or low value (or a middle value) when the other variable is small (large), or vice versa.

One example is the relationship between access to social media and press freedom. Kocak and Kıbrıs (2023) develop a game-theoretic model to argue that higher internet access promotes press freedom in a country when an incumbent has a lower rent from office but prohibits it in a country with higher rent. The left panel of Figure 1 replicates Figure 1 of Kocak and Kıbrıs (2023) that shows the relationship between internet penetration and press freedom in 160 countries from 2000 to 2015.[3] Each dot corresponds to a unit of observation, namely, a country in a year ($n = 2,560$). The horizontal axis represents internet penetration. The vertical axis indicates a press freedom score. When internet penetration is low, the press freedom scores vary (near the left vertical axis). When internet penetration is high, some units have high press freedom scores (top-right corner) and others have low press freedom scores (bottom-right corner), although no unit has a moderate press freedom score (near and in the middle of the right vertical axis). Therefore, the relationship between the two variables is nonmonotonic in the sense that internet penetration decreases in press freedom when press freedom is low but increases in press freedom when press freedom is high. Since the rent is unobserved, we cannot condition on it and should have studied the joint distribution.

Another instance is civil war duration and outcome, 1946 to 2003 (Cunningham, Gleditsch, and Salehyan 2009). In the right panel of Figure 1, each point represents a civil war ($n = 267$), the horizontal axis indicates the duration in days, and the vertical axis corresponds to the outcome, which is equal

[3]I obtained the dataset from one of the authors, Korhan Kocak, through personal communication on March 23, 2023. The replication materials for this article can be found at Fukumoto (2023b).

to one if the government wins, two if neither side wins, and three if the rebellion wins.[4] If either the government or the rebellion is strong enough to win, the civil war ends in a short time. Otherwise, neither side obtains a decisive victory, but they instead reach a negotiated settlement to spare the continuing war attrition costs. They do this only after the intrastate military conflict persists for a long duration (Mason, Weingarten Jr, and Fett 1999). Here, the relative power of rebels to governments is a confounder, which is observed but only as a three-category indicator with measurement error.

In these cases, the underlying copulas have irregular properties, such as nonmonotonicity and asymmetricity. Few parametric copulas can handle these properties. To address this gap, I pay special attention to an overlooked copula (Chesneau 2021), which I name the "normal mode copula." It is imperative to enlarge the pool of copulas so that analysts can flexibly adapt the joint distribution to the data and alleviate reliance on the functional-form assumption of the joint distribution (Braumoeller *et al.* 2018).

This article is organized as follows: The next section elaborates on the definition of copulas, introduces some conventional copulas, and defines the normal mode copula. In the following section, I apply these copulas to a dataset on government formation and duration (Chiba *et al.* 2015) to demonstrate that the normal mode copula fits the data better than do other conventional copulas. Finally, I present concluding remarks.

2. Definition

2.1. Generic and Gaussian Copulas

Let $0 \le u_d \le 1$ for $d \in \{1,2\}$. A function $C(u_1,u_2):[0,1]^2 \to [0,1]$ is called a copula if the following two conditions are met (Nelsen 2006, 10):

> ***Boundary conditions***: $C(u_1,0)=C(0,u_2)=0, C(u_1,1)=u_1$, and $C(1,u_2)=u_2$.
> ***Two-increasing condition***: If $u_1^L \le u_1^H$ and $u_2^L \le u_2^H$, it follows that $C(u_1^H,u_2^H)-C(u_1^L,u_2^H)-C(u_1^H,u_2^L)+C(u_1^L,u_2^L) \ge 0$.

The motivation for copulas is as follows. I suppose that there are two random variables, X_1 and X_2. I denote the value of the dth variable X_d by x_d, the marginal cumulative distribution function (CDF) of X_d by $F_d(x_d) \equiv u_d$, and the joint CDF of X_1 and X_2 by $F_{12}(x_1,x_2)$. Then, according to Sklar's theorem (Nelsen 2006, 18 and 24–25), there exists a copula C such that

$$F_{12}(x_1,x_2) = C\{F_1(x_1),F_2(x_2)\}. \tag{1}$$

When F_1 and F_2 are continuous, C is unique. If X_1 and X_2 are continuous variables and we can differentiate both sides of Equation (1) by x_1 and x_2, we obtain

$$f_{12}(x_1,x_2) = f_1(x_1)f_2(x_2)c(u_1,u_2), \tag{2}$$

where $f_d(x_d)$ is the probability density function (PDF) of X_d and

$$c(u_1,u_2) \equiv \frac{\partial^2}{\partial u_1 \partial u_2} C(u_1,u_2)$$

represents densities for copula C.[5] The copula C of X_1 and X_2 abstracts away the marginal distributions of X_1 and X_2 and thus distills all information about the dependence between X_1 and X_2. Equation (2) clarifies the modularity of the copula: we can substitute the marginal distribution of one variable

[4] I downloaded the replication materials of Cunningham *et al.* (2009) (the original publication version, not the latest updated version) from http://jcr.sagepub.com/content/53/4/570/suppl/DC1 on July 6, 2012. I preprocessed the dataset in the same way as Fukumoto (2015) did.

[5] In this case, the two-increasing condition is equivalent to $c(u_1,u_2) \ge 0$.

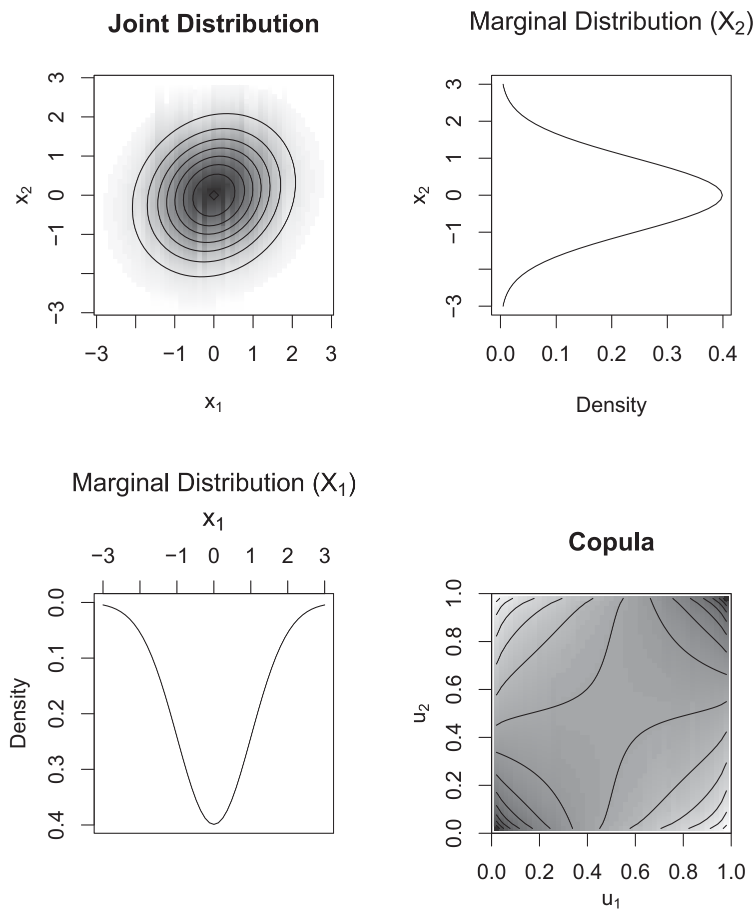

Figure 2. Bivariate normal distribution: joint distribution, marginal distributions, and copula.

or the copula without changing the other two terms on the right-hand side to obtain a new bivariate distribution on the left-hand side.

For instance, the Gaussian copula is defined as

$$C_G(u_1,u_2) \equiv \Phi^2\{\Phi^{-1}(u_1),\Phi^{-1}(u_2) \mid \theta\},$$

where Φ and Φ^2 are univariate and bivariate standard normal distributions, respectively, and $-1 \leq \theta \leq 1$ is the correlation parameter. Figure 2 illustrates an example; the top-left panel shows a contour plot of densities for the joint distribution $f_{12}(x_1,x_2) = \phi^2(x_1,x_2 \mid \theta = 0.156)$, where ϕ^2 is the PDF of Φ^2;[6] the bottom-left and top-right panels present densities for the marginal distributions $f_1(x_1) = \phi(x_1)$ and $f_2(x_2) = \phi(x_2)$, respectively, where ϕ is the PDF of Φ; the bottom-right panel represents a contour plot of densities for the copula $c(u_1,u_2)$, which turns to be Gaussian, $c_G(u_1,u_2)$.[7] If we substitute $f_1(x_1)$,$f_2(x_2)$, or $c(u_1,u_2)$, the joint distribution $f_{12}(x_1,x_2)$ is no longer represented by ϕ^2.

[6]In Figures 2 and 3, the value of θ is set so that the Kendall rank correlation (Nelsen 2006, 161 and 164, Trivedi and Zimmer 2007, 16 and 22), a measure of association, is equal to 0.1, and thus copulas are comparable. I conduct all analyses in the statistical computational environment `R` (R Core Team 2022). For Gaussian copulas, I employ the `mvtnorm` library (Genz *et al.* 2021).

[7]In Figures 2–6, the shade in the contour plot is decided according to the maximum densities of each panel and thus is not comparable across panels.

2.2. Conventional Copulas

Scholars have derived dozens of copulas (Hofert *et al.* 2018; Nelsen 2006; Trivedi and Zimmer 2007). For reference, I introduce the product copula and the following four conventional copulas: Farlie–Gumbel–Morgenstein (FGM), Ali–Mikhail–Haq (AMH), Clayton, and Frank.

The product (or independence) copula is defined as (Nelsen 2006, 25)

$$C_{\text{I}}(u_1,u_2) \equiv u_1 u_2,$$

where X_1 and X_2 are independent of each other.

The FGM copula is defined as (Nelsen 2006, 77)

$$C_{\text{FGM}}(u_1,u_2) \equiv u_1 u_2 + \theta u_1(1-u_1)u_2(1-u_2),$$

where $-1 \leq \theta \leq 1$.

The other three belong to a class of Archimedean copulas. Let the generator function $\varphi(u) : [0,1] \to [0,\infty]$ be a continuous, convex, strictly decreasing function, where $\varphi(1) = 0$. I define the pseudoinverse function of $\varphi(u)$ as $\varphi^{[-1]}(z) = \varphi^{-1}(z)$ for $0 \leq z \leq \varphi(0)$ and $\varphi^{[-1]}(z) = 0$ for $\varphi(0) \leq z \leq \infty$. Then, the following function:

$$C_{\text{A}}(u_1,u_2 \mid \varphi) \equiv \varphi^{[-1]}\{\varphi(u_1)+\varphi(u_2)\}$$

meets the two conditions of a copula and is called an Archimedean copula (Nelsen 2006, 110–112). The generator functions for the AMH, Clayton, and Frank copulas are

$$\begin{aligned}
\varphi_{\text{AMH}}(u) &\equiv \log\left[\frac{1}{u}\{1-\theta(1-u)\}\right] && \text{for} \quad -1 \leq \theta \leq 1,\\
\varphi_{\text{Clayton}}(u) &\equiv \theta^{-1}(u^{-\theta}-1) && \text{for} \quad \theta > 0,\\
\varphi_{\text{Frank}}(u) &\equiv -\log\left[\frac{1}{\exp(-\theta)-1}\{\exp(-\theta u)-1\}\right] && \text{for} \quad \theta \in \mathbb{R},
\end{aligned}$$

respectively.

Figure 3 shows contour plots of densities for the conventional copulas, where the horizontal and vertical axes are u_1 and u_2, respectively. Panels correspond to FGM, AMH, Clayton, and Frank.[8] Clearly, these copulas represent monotonic dependence.

In general, the three associated copulas of a copula are defined as (Trivedi and Zimmer 2007, 13–14)

$$\begin{aligned}
\overline{C}^{(1)}(u_1,u_2) &\equiv u_2 - C(1-u_1,u_2),\\
\overline{C}^{(2)}(u_1,u_2) &\equiv u_1 - C(u_1,1-u_2),\\
\overline{C}^{(12)}(u_1,u_2) &\equiv u_1+u_2-1+C(1-u_1,1-u_2).
\end{aligned} \tag{3}$$

In particular, $\overline{C}^{(12)}(u_1,u_2)$ is called the survival copula (Nelsen 2006, 32). Figure 4 illustrates contour plots of densities for the associated Clayton copulas. Graphically, by turning the density $c(u_1,u_2)$ (first panel, the same as the bottom-left panel of Figure 3) with respect to the line $u_d = \frac{1}{2}$, we obtain $\overline{c}^{(d)}(u_1,u_2)$ (second ($d=1$) and third ($d=2$) panels); by rotating $c(u_1,u_2)$ 180 degrees, we obtain $\overline{c}^{(12)}(u_1,u_2)$ (fourth panel).[9] Specifically, I study the three associated copulas for AMH and Clayton copulas (cf. Braumoeller *et al.* 2018, 59); for each of the other conventional copulas, the three associated copulas belong to the family of the original copula.

[8]The values of θ are equal to 0.45, 0.381, 0.222, and 0.907 in the order of the mentioned copulas, respectively. For deriving θ of the Frank copula, I utilize the `gsl` library (version 2.1-7.1 Hankin 2006).

[9]The Kendall rank correlation is equal to 0.1, -0.1, -0.1, and 0.1 for $c, \overline{c}^{(1)}, \overline{c}^{(2)}$, and $\overline{c}^{(12)}$, respectively.

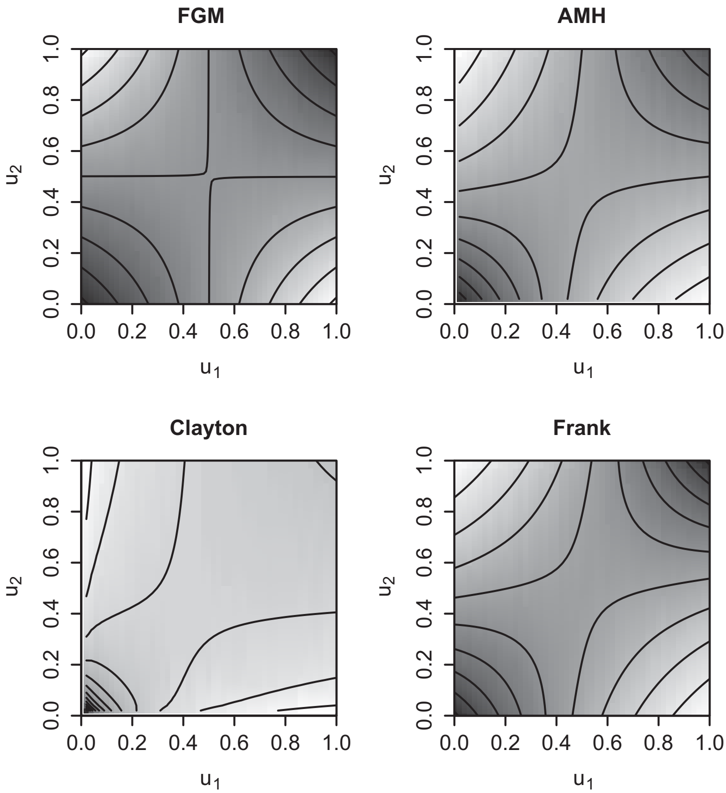

Figure 3. Example plots of conventional copulas.

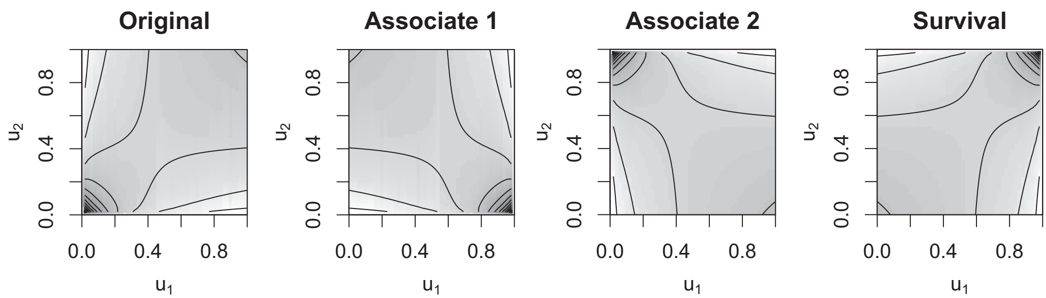

Figure 4. Example plots of associated Clayton copulas.

2.3. *Normal Mode Copulas*

Chesneau (2021) refers to the following copula but only in passing:

$$C_{\text{Chesneau}}(u_1,u_2) = u_1u_2 + \frac{1}{p_1p_2\kappa_1\kappa_2\pi^2}\theta\{\sin(u_1\kappa_1\pi)\}^{p_1}\{\sin(u_2\kappa_2\pi)\}^{p_2}, \tag{4}$$

where for $d \in \{1,2\}$, κ_d's are positive integers, $p_d \geq 1$, and $-1 \leq \theta \leq 1$.[10] However, Chesneau (2021) did not apply the copula to any real data or give it any name.

[10]To be accurate, Chesneau (2021) introduces the multivariate version of this copula.

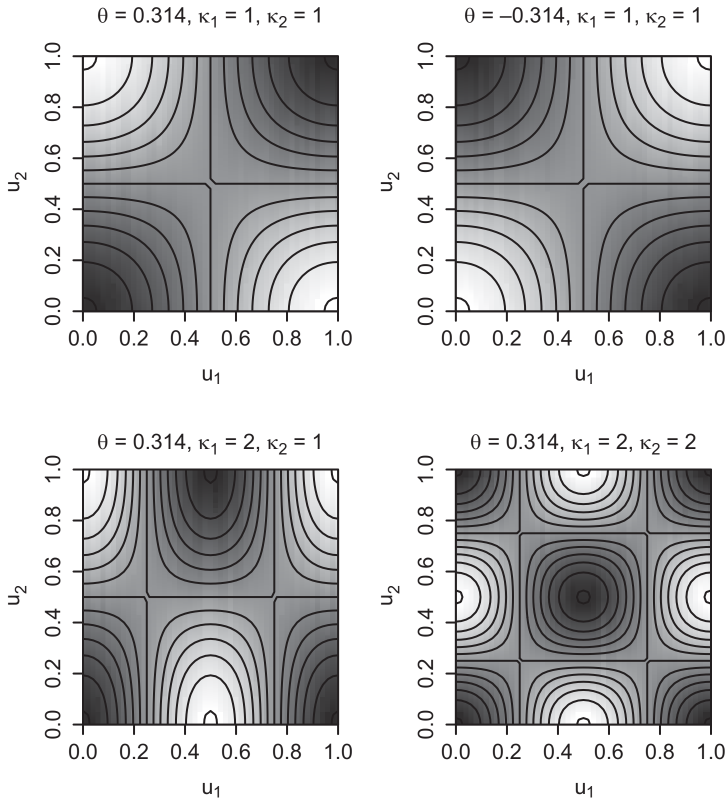

Figure 5. Example plots of normal mode copulas.

I simplify this copula by setting $p_d = 1$,

$$C_{\text{NM}}(u_1,u_2) \equiv u_1u_2 + \frac{1}{\kappa_1\kappa_2\pi^2}\theta\sin(u_1\kappa_1\pi)\sin(u_2\kappa_2\pi),$$

so that the copula is still flexible in a tractable way with fewer parameters. I name C_{NM} the normal mode copula, θ the amplitude, and κ_d's mode numbers. (I explain the motivation shortly.)

Figure 5 shows contour plots of densities for normal mode copulas,

$$c_{\text{NM}}(u_1,u_2) = 1+\theta\cos(u_1\kappa_1\pi)\cos(u_2\kappa_2\pi).$$

The top-left panel corresponds to the case where $\theta = 0.304, \kappa_1 = 1$, and $\kappa_2 = 1$.[11] This normal mode copula is (radially) symmetric, where U_2 monotonically increases in U_1. In the top-right panel, I present another normal mode copula where I only change θ from 0.304 to -0.304. We obtain the current plot by making the dark parts in the previous plot light and vice versa. Here, U_2 monotonically decreases in U_1. The bottom-left panel represents the case where $\theta = 0.304, \kappa_1 = 2$, and $\kappa_2 = 1$. This normal mode copula is *not* (radially) symmetric. Moreover, U_2 monotonically increases in U_1 for $U_1 \le \frac{1}{2}$ but decreases in U_1 for $U_1 \ge \frac{1}{2}$. Thus, the copula represents a case of nonmonotonic dependence that motivates me, as I explained in Section 1. In fact, if we rotate the plot clockwise (counterclockwise) 90 degrees, it resembles the right (left) panel of Figure 1. Finally, the bottom-right panel addresses the case where

[11]The Kendall rank correlation is equal to $0.1, -0.1, 0$, and 0 for the first to fourth panels, respectively. Note that the Kendall rank correlation can be equal to 0 when the two variables are not independent of each other and the dependence between these two variables is nonmonotonic.

$\theta = 0.304, \kappa_1 = 2$, and $\kappa_2 = 2$. This normal mode copula is (radially) symmetric but has nonmonotonic dependence.

If we regard u_1 as the position along an open pipe (such as a flute) with unit length and u_2 as time, the displacement of a standing sound wave with one frequency at position u_1 and time u_2 resembles $c_{NM}(u_1, u_2)$.[12] Importantly, if κ_1 were not an integer, the corresponding sound wave would not form a standing wave and would not resonate, thus dissipating immediately. In general, this kind of standing wave is called the normal mode. This is why I name C_{NM} the normal mode copula.

3. Application

3.1. Overview

This section intends to showcase the usefulness of normal mode copulas by reanalyzing Chiba *et al.* (2015, hereafter, "CMS"), who also use a copula.[13] CMS focus on two outcomes of (coalition) government formation and duration. Scholars have studied what factors affect either outcome. One problem is that these two outcomes are interrelated, and failure to consider simultaneity leads to selection bias in estimating the model. For instance, if a government would not remain in power for a long time, say, due to a hidden scandal, it may be less likely to be formed in the first place. If (and probably as) scholars do not observe all confounders (e.g., the hidden scandal) that affect both outcomes or do observe some of them but with measurement error, the resultant estimates will be less efficient and might suffer from omitted variable bias. To address this problem, CMS incorporate a copula into their model and analyze the two outcomes jointly. A remaining concern is that CMS consider only the Gaussian copula. I substitute the normal mode copulas as well as other conventional copulas to show that a normal mode copula is the best.

3.2. Model

To focus on the comparison between the Gaussian copula and other copulas, this subsection introduces a simplified version of the CMS model using my notation.[14] The units of observation are a (coalition government) formation opportunity $i \in \{1, 2, \ldots, n = 432\}$ and a potential coalition government (combination of parties) $j \in \mathcal{J}_i \equiv \{1, 2, \ldots, m_i\}$ ($\sum_i m_i = 95,576$). For each i, we have two outcome variables. The first outcome is a realized government, $Y_{1,i} \in \mathcal{J}_i$. The second outcome is the duration of the government from its inception to its termination, $Y_{2,i} \in (0, \bar{y}_i]$, where $\bar{y}_i$ is the constitutional interelection period. Unless a government ends for political reasons (replacement by another government without an election or dissolution of the legislature followed by an early election), I regard the government's duration as censored.[15]

CMS use a conditional logit model to explain government formation. We denote a covariate vector by $\boldsymbol{z}_{1,ij}$ and the corresponding coefficient vector by $\boldsymbol{\theta}_1$. Specifically, $\boldsymbol{z}_{1,ij}$ is composed of Minority Government, Status Quo Government, and dozens of party dummy variables.[16] The probability that

[12] Strictly speaking, the displacement is $\theta \cos(u_1 \kappa_1 \pi) \cos(u_2 \kappa_2 \pi)$, where $|\theta|$ can be larger than one, X_2 is a circular variable, and κ_2 is an even integer.

[13] I downloaded the replication materials (Chiba, Martin, and Stevenson 2014) on March 14, 2023. The dataset covers 17 European countries, 1945 to 2011. For details, see CMS, f.n. 11 (p.51).

[14] I thank a reviewer for suggesting simplifying the CMS model. In the Supplementary Material, I present reanalyses of the full CMS model.

[15] According to CMS, "an observation is right-censored if the government was still in office as of December 31, 2011 (the cutoff date for our sample), or if it terminated for any of the following reasons: the occurrence of a regularly scheduled election, a technical resignation required for constitutional reasons, or the death of the prime minister. . . . Of our 432 governments, 89 are right-censored for these reasons" (p. 52). CMS employ the competing risks approach where they differentiate termination due to replacement from termination due to dissolution.

[16] Minority is equal to one if government parties do not collectively control a parliamentary majority, zero otherwise (CMS, 53). Status Quo Government is equal to one if the potential coalition government is the same as the pervious government,

government $g \in \mathcal{J}_i$ is formed is modeled as

$$\Pr(Y_{1,i} = g) = \frac{\exp(\boldsymbol{z}_{1,ig}'\boldsymbol{\theta}_1)}{\sum_{j\in\mathcal{J}_i}\exp(\boldsymbol{z}_{1,ij}'\boldsymbol{\theta}_1)} \equiv G(g \mid \boldsymbol{z}_{1,i}, \boldsymbol{\theta}_1),$$

where $\boldsymbol{z}_{1,i} \equiv (\boldsymbol{z}_{1,i1}', \boldsymbol{z}_{1,i2}', \ldots, \boldsymbol{z}_{1,im_i}')'$. In my understanding, CMS implicitly assume that when a latent utility variable $X_{1,i}$ is smaller than $\bar{x}_{1,i} \equiv F_1^{-1}\{G(g \mid \boldsymbol{z}_{1,i}, \boldsymbol{\theta}_1)\}$, we observe $Y_{1,i} = g$.[17]

CMS's main model assumes that the government duration follows a Weibull distribution, $F_2(t \mid \theta_{2,\text{inv.scale}}, \theta_{2,\text{shape}})$, where $\theta_{2,\text{inv.scale}}$ is the inverse scale parameter and $\theta_{2,\text{shape}}$ is the shape parameter. We denote the latent time variable by $X_{2,i}$, another covariate vector by $\boldsymbol{z}_{2,i}$, and the corresponding coefficient vector by $\boldsymbol{\theta}_{2,\text{coef}}$. To be concrete, I include Minority and Polarization Index in $\boldsymbol{z}_{2,i}$.[18] The probability density that the length of the duration is t is modeled as

$$p(X_{2,i} = t) = f_2(t \mid \theta_{2,\text{inv.scale}}, \theta_{2,\text{shape}}),$$

where $\theta_{2,\text{inv.scale}} = \exp(-\boldsymbol{z}_{2,i}'\boldsymbol{\theta}_{2,\text{coef}})$.

Let W_i be the censoring indicator. If the duration ends for political reasons at $t < \bar{y}_i$, we observe $W_i = 0$, $Y_{1,i} = g$, and $Y_{2,i} = X_{2,i} = t$, and the mixed joint density is

$$\begin{aligned} p(Y_{1,i} = g, Y_{2,i} = t) &= \Pr(X_{1,i} \le \bar{x}_{1,i} \mid X_{2,i} = t)p(X_{2,i} = t) \\ &= F_{1|2}(\bar{x}_{1,i} \mid X_{2,i} = t)f_2(t \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2) \\ &= C_{1|2}'(\bar{u}_{1,i} \mid \underline{u}_{2,i}, \theta_{12})f_2(t \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2), \end{aligned} \tag{5}$$

where $\bar{u}_{1,i} = F_1(\bar{x}_{1,i}) = G(g \mid \boldsymbol{z}_{1,i}, \boldsymbol{\theta}_1)$, $\underline{u}_{2,i} \equiv F_2(t \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)$, $\boldsymbol{\theta}_2 \equiv (\boldsymbol{\theta}_{2,\text{coef}}', \theta_{2,\text{shape}})'$, θ_{12} is the parameter of copula C of X_1 and X_2, and it generally holds that $\Pr(X_1 \le x_1 \mid X_2 = x_2) = \frac{\partial}{\partial u_2}C(u_1, u_2) \equiv C_{1|2}'(u_1 \mid u_2)$ (Nelsen 2006, 41).

If the duration is censored at t, we observe $W_i = 1$, $Y_{1,i} = g$, and $Y_{2,i} = t$ but not $X_{2,i} > t$, and the joint probability is

$$\begin{aligned} \Pr(Y_{1,i} = g, Y_{2,i} = t) &= \Pr(X_{1,i} \le \bar{x}_{1,i}, X_{2,i} > t) \\ &= \Pr(X_{1,i} \le \bar{x}_{1,i}) - \Pr(X_{1,i} \le \bar{x}_{1,i}, X_{2,i} \le t) \\ &= F_1(\bar{x}_{1,i}) - F_{12}(\bar{x}_{1,i}, t) \\ &= \bar{u}_{1,i} - C(\bar{u}_{1,i}, \underline{u}_{2,i} \mid \theta_{12}). \end{aligned} \tag{6}$$

By multiplying either Equation (5) or (6), we can obtain the total likelihood function,

$$\begin{aligned} \mathcal{L}_{12}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \theta_{12} \mid \boldsymbol{w}, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{Z}_1, \boldsymbol{Z}_2) \propto \prod_{i=1}^{n} &\left\{C_{1|2}'(\bar{u}_{1,i} \mid \underline{u}_{2,i}, \theta_{12})f_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)\right\}^{I(w_i=0)} \\ &\left\{\bar{u}_{1,i} - C(\bar{u}_{1,i}, \underline{u}_{2,i} \mid \theta_{12})\right\}^{I(w_i=1)}, \end{aligned} \tag{7}$$

where $I(\cdot)$ is the dummy indicator function, and maximize it to estimate the parameters, $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$, and θ_{12}. This is the CMS model.

---

zero otherwise (Martin and Stevenson 2010, 504). The intercept is not included in $\boldsymbol{z}_{1,ij}$. CMS include 20 additional variables. I retain Minority because CMS show that it affects both formation and duration in a statistically significant manner. I retain Status Quo Government for the exclusion restriction; CMS (53–54) argue that it affects formation but not duration. See also the Supplementary Material. As for party dummy variables, see Martin and Stevenson (2010, 512–514).

[17] For details, see the Supplementary Material. In other words, $X_{1,i}$ is integrated out: $\Pr(Y_{1,i} = g) = \int_{-\infty}^{\bar{x}_{1,i}} f_1(x_1)dx_1$.

[18] Polarization Index measures the presence of anti-establishment parties in the legislature (for details, see CMS, 53). The intercept is included in $\boldsymbol{z}_{2,i}$. CMS add four variables. I retain Polarization Index for the purpose of the exclusion restriction; CMS (53–54) argue that it affects duration but not formation.

3.3. Studied Copulas

CMS assume that the copula of X_1 and X_2 is a Gaussian copula. My departure from the CMS model starts here. I substitute normal mode copulas. In general, before analysts apply a normal mode copula to any dataset, they have to determine the values of κ_1 and κ_2, which are model choice indicators rather than parameters to be estimated. In my case, I explore $\kappa_1, \kappa_2 \in \{1,2,3,4\}$. I also consider the following four conventional copulas: FGM, AMH, Clayton, and Frank. The three associated copulas of AMH and Clayton copulas are also studied.

For the purpose of comparison, I analyze the "separate" model of formation and duration by maximizing each of the following likelihood functions:

$$\mathcal{L}_1(\boldsymbol{\theta}_1 \mid \boldsymbol{y}_1, \boldsymbol{Z}_1) \propto \prod_{i=1}^{n} G(y_{1,i} \mid \boldsymbol{z}_{1,i}, \boldsymbol{\theta}_1), \tag{8}$$

$$\mathcal{L}_2(\boldsymbol{\theta}_2 \mid \boldsymbol{w}, \boldsymbol{y}_2, \boldsymbol{Z}_2) \propto \prod_{i=1}^{n} \{f_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)\}^{I(w_i=0)} \{1 - F_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)\}^{I(w_i=1)}. \tag{9}$$

I also analyze another duration model where I include the predicted probability of formation ($\hat{\overline{u}}_{1,i}$, hereafter, "Formation Probability"), which is estimated by using the formation model (Equation (8)), as a covariate:

$$\mathcal{L}_2^*(\boldsymbol{\theta}_2^* \mid \boldsymbol{w}, \boldsymbol{y}_2, \boldsymbol{Z}_2, \hat{\overline{\boldsymbol{u}}}_1) \propto \prod_{i=1}^{n} \left\{f_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \hat{\overline{u}}_{1,i}, \boldsymbol{\theta}_2^*)\right\}^{I(w_i=0)} \left\{1 - F_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \hat{\overline{u}}_{1,i}, \boldsymbol{\theta}_2^*)\right\}^{I(w_i=1)}. \tag{10}$$

If the relationship between formation and duration is monotonic, this "two-step" model should work.[19] Note that the separate and two-step models are effectively equivalent to the CMS model with the product copula (C_I), which has no parameter, because $C'_{I,1|2}(u_1 \mid u_2) = u_1$ and Equation (7) becomes

$$\begin{aligned}
&\prod_{i=1}^{n} \{C'_{I,1|2}(\overline{u}_{1,i} \mid \underline{u}_{2,i}) f_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)\}^{I(w_i=0)} \{\overline{u}_{1,i} - C_I(\overline{u}_{1,i}, \underline{u}_{2,i})\}^{I(w_i=1)} \\
&= \prod_{i=1}^{n} \{\overline{u}_{1,i} f_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)\}^{I(w_i=0)} \{\overline{u}_{1,i} - \overline{u}_{1,i}\underline{u}_{2,i}\}^{I(w_i=1)} \\
&= \prod_{i=1}^{n} \underbrace{G(y_{1,i} \mid \boldsymbol{z}_{1,i}, \boldsymbol{\theta}_1)}_{\text{formation}} \underbrace{\{f_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)\}^{I(w_i=0)} \{1 - F_2(y_{2,i} \mid \boldsymbol{z}_{2,i}, \boldsymbol{\theta}_2)\}^{I(w_i=1)}}_{\text{duration}}.
\end{aligned}$$

If copula models improve the model fit, we can alleviate bias due to unobserved confounders and root mean squared error (CMS, Braumoeller *et al.* 2018). I also expect that a copula model that fits the data better leads to smaller standard errors because the copula model takes greater advantage of the information about formation (X_1) in estimating the parameters about duration (X_2), and thus the conditional variance of duration given formation is smaller.

3.4. Results

Table 1 reports the Akaike information criterion (AIC) for each copula used in the CMS model.[20] Each row indicates the model with each copula. In the first and second rows, I report the AICs of the separate and two-step models, respectively, where I sum the AICs of the formation model (Equation (8)) and the duration model (Equation (9) for the separate model and Equation (10) for the two-step model). The AIC of the two-step model is larger than that of the separate model almost by two, which means the predicted probability of formation has little linear relation with duration. Below, all models employ Equation (7). Compared with the separate model, the Gaussian copula model CMS used (third row)

[19] I thank a reviewer for suggesting the two-step model.
[20] Braumoeller *et al.* (2018) and Fukumoto (2015) also select the best among copulas in terms of AIC.

Table 1. AICs for models using various copulas.

Copula		AIC
Product	Separate	7,686.2
	Two-step	7,688.2
Gaussian		7,686.7
Normal Mode	$\kappa_1 = 1, \kappa_2 = 1$	7,686.3
	$\kappa_1 = 1, \kappa_2 = 2$	7,628.2
	$\kappa_1 = 1, \kappa_2 = 3$	7,644.1
	$\kappa_1 = 1, \kappa_2 = 4$	7,644.3
	$\kappa_1 = 2, \kappa_2 = 1$	7,649.0
	$\kappa_1 = 2, \kappa_2 = 2$	7,668.8
	$\kappa_1 = 2, \kappa_2 = 3$	7,688.1
	$\kappa_1 = 2, \kappa_2 = 4$	7,674.0
	$\kappa_1 = 3, \kappa_2 = 1$	7,677.5
	$\kappa_1 = 3, \kappa_2 = 2$	7,638.7
	$\kappa_1 = 3, \kappa_2 = 3$	7,661.5
	$\kappa_1 = 3, \kappa_2 = 4$	7,646.5
	$\kappa_1 = 4, \kappa_2 = 1$	7,638.0
	$\kappa_1 = 4, \kappa_2 = 2$	7,685.6
	$\kappa_1 = 4, \kappa_2 = 3$	7,668.9
	$\kappa_1 = 4, \kappa_2 = 4$	7,680.9
FGM		7,688.0
AMH	Original	7,687.8
	Associate 1	7,688.0
	Associate 2	7,687.9
	Survival	7,688.0
Clayton	Original	7,688.1
	Associate 1	7,688.2
	Associate 2	7,688.2
	Survival	7,688.2
Frank		7,687.9

worsens the AIC. In the next 16 rows, I display the AICs of normal mode copula models. The normal mode copula model with $\kappa_1 = 1$ and $\kappa_2 = 2$ (hereafter, the "NM(1, 2) copula model," fifth row) has the best (i.e., smallest) AIC. In the last ten rows, I present the AICs of the conventional copula models and their associated copula models. (For the AMH and Clayton copulas, "Original," "Associate 1," "Associate 2," and "Survival" indicate C, $\overline{C}^{(1)}$, $\overline{C}^{(2)}$, and $\overline{C}^{(12)}$ (Equation (3)), respectively.) They have almost the same AIC as the two-step model and do not outperform the NM(1, 2) copula model. This is probably because all of these conventional copulas cope with monotonic relationships alone, where in fact, the relation between $\overline{u}_{1,i}$ and $\underline{u}_{2,i}$ is nonmonotonic as shown next.

Figure 6 illustrates the scatter plots of estimated $\overline{u}_{1,i}$'s (horizontal axis) and $\underline{u}_{2,i}$'s (vertical axis). The left and right panels correspond to the Gaussian copula model and the NM(1, 2) copula model,

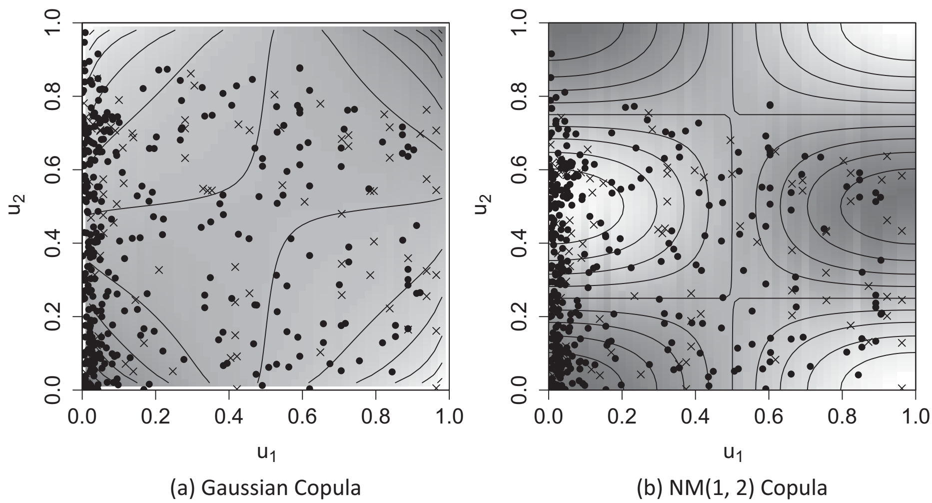

Figure 6. Scatter plots of $\overline{u}_{1,i}$'s and $\underline{u}_{2,i}$'s with contour plots of the estimated copula. $n = 432$.

respectively. In each panel, the contour plot of the copula based on the estimate of θ_{12} is overlaid.[21] A point indicates a unit that is not censored, where $(U_{1,i}, U_{2,i}) \in \{(u_1, u_2) \mid 0 < u_1 \leq \overline{u}_{1,i}, u_2 = \underline{u}_{2,i}\}$. A cross indicates a unit that is censored, where $(U_{1,i}, U_{2,i}) \in \{(u_1, u_2) \mid 0 < u_1 \leq \overline{u}_{1,i}, \underline{u}_{2,i} < u_2 < 1\}$. Clearly, the relationship between $\overline{u}_{1,i}$ and $\underline{u}_{2,i}$ is nonmonotonic. Since I use a simplified CMS model, neither copula model appears to fit the data well. However, units near the right vertical axis are situated in low-density areas of the Gaussian copula model but in high-density areas of the NM(1, 2) copula model. This is likely why the NM(1, 2) copula model achieves the best performance among the studied copulas in Table 1.

Table 2 presents the estimation results of the parameters.[22] The first two rows display covariate coefficients of the formation model ($\boldsymbol{\theta}_1$ except for dozens of party fixed effects). The third to sixth rows indicate covariate coefficients of the duration model (including intercept, $\boldsymbol{\theta}_{2,\text{coef}}$), while the last row represents the logged shape parameter of the duration model ($\log(\theta_{\text{shape}})$). The first and second columns concern the separate model (Equations (8) and (9), respectively), while the third and fourth columns correspond to the two-step model (Equations (8) and (10), respectively; thus, the first column is the same as the third column). In the fifth and sixth columns, I show the results of the Gaussian copula model and the NM(1, 2) copula model, respectively (Equation (7)). In each cell, the entry is the estimate with the standard error in parentheses.

As expected, all standard errors are smaller in the NM(1, 2) copula model than in the separate model, the two-step model, and the Gaussian copula model, except for that of the Duration Dependence parameter. In the case of Minority Government as a formation covariate, the standard error of the NM(1, 2) copula model is reduced by 34% compared with that of the separate model. The coefficient of Formation Probability in the two-step model is close to zero and not significant, although the NM(1, 2) copula model performs well. (Recall also that the two-step model does not improve the AIC compared with the separate model in Table 1.) This implies that the probability of formation affects duration not in a monotonic way but in a nonmonotonic way. Point estimates are not particularly different across models in the simplified models. However, if I use the full CMS model, they are so distinct that coefficient estimates are statistically significantly different from zero in one model but not in another (Supplementary Material).

[21] $\hat{\theta}_{12} = 0.070$ for the Gaussian copula and $\hat{\theta}_{12} = 1.000$ for the NM(1, 2) copula. The corresponding Kendall rank correlations are 0.045 and 0, respectively, which implies little linear correlation between formation and duration.

[22] Chiba *et al.* (2014) require us to use the `corpcor` library (Schafer *et al.* 2021).

Table 2. Results of parameter estimation.

	Separate		Two-step		Gaussian	NM(1, 2)
Formation						
Minority Government	−1.370		−1.370		−1.384	−1.341
	(0.180)		(0.180)		(0.121)	(0.119)
Status Quo Government	2.614		2.614		2.625	2.490
	(0.125)		(0.125)		(0.119)	(0.118)
Duration						
Minority Government		−0.390		−0.388	−0.369	−0.389
		(0.085)		(0.086)	(0.084)	(0.079)
Polarization Index		−0.068		−0.067	−0.065	−0.059
		(0.020)		(0.020)	(0.020)	(0.018)
Formation Probability				0.018		
				(0.143)		
Intercept		7.059		7.054	7.138	7.380
		(0.055)		(0.070)	(0.054)	(0.053)
Duration Dependence		0.305		0.305	0.347	0.299
		(0.045)		(0.045)	(0.045)	(0.045)

Note: Cell entries are estimates (with standard errors in parentheses). $n = 432$.

4. Conclusion

This study sheds light on an understudied family of copulas, normal mode copulas. I apply dozens of copulas to a dataset of government formation and duration to show that the normal mode copula achieves higher performance than other conventional copulas.

There are several directions for future research. In a companion paper (Fukumoto 2023a), I have characterized the properties of normal mode copulas such as monotonicity and measures of association. It is also promising to explore the properties of Chesneau's (2021) copulas (Equation (4)) and its multivariate version. Scholars can apply normal mode copulas to various data to find interesting dependence structures. I hope normal mode copulas become a helpful tool to analyze mutually dependent variables in political science.

Acknowledgments. Earlier versions of this article were presented on November 5, 2013, June 23, 2015, and September 16, 2022, at the Institute of Statistical Mathematics, Tokyo, Japan, and on January 6–7, 2023, at the 10th Asian Political Methodology Meeting, National University of Singapore. I thank Toshikazu Kitano and Dean Knox for their helpful comments. I am grateful to Korhan Kocak, Özgür Kıbrıs, David E. Cunningham, Kristian Skrede Gleditsch, Idean Salehyan, Daina Chiba, Lanny W. Martin, and Randolph T. Stevenson for enabling me to reanalyze their datasets. I also appreciate American Journal Experts for providing proofreading service. I am entirely responsible for the scientific content of the article, and the article adheres to the journal's authorship policy.

Funding. This work was supported by the Japan Society for the Promotion of Science (Grant No. KAKENHI JP19K21683) and the Gakushuin University's Computer Centre (no grant number).

Conflicts of Interest. There are no conflicts of interest to disclose.

Data Availability Statement. The replication materials for this article can be found in Fukumoto (2023b) at https://doi.org/10.7910/DVN/X94ITA.

Supplementary Material. For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2023.45.

References

Braumoeller, B. F., G. Marra, R. Radice, and A. E. Bradshaw. 2018. "Flexible Causal Inference for Political Science." *Political Analysis* 26 (1): 54–71.

Chesneau, C. 2021. "On New Types of Multivariate Trigonometric Copulas." *Applied Mathematics* 1 (1): 3–17.

Chiba, D., L. W. Martin, and R. T. Stevenson. 2014. Replication Data for: A Copula Approach to the Problem of Selection Bias in Models of Government Survival. https://doi.org/10.7910/DVN/26966, Harvard Dataverse, V2.

Chiba, D., L. W. Martin, and R. T. Stevenson. 2015. "A Copula Approach to the Problem of Selection Bias in Models of Government Survival." *Political Analysis* 23 (1): 42–58.

Chiba, D., N. W. Metternich, and M. D. Ward. 2015. "Every Story Has a Beginning, Middle, and an End (But Not Always in That Order): Predicting Duration Dynamics in a Unified Framework." *Political Science Research and Methods* 3 (3): 515–541.

Cunningham, D. E., K. S. Gleditsch, and I. Salehyan. 2009. "It Takes Two: A Dyadic Analysis of Civil War Duration and Outcome." *Journal of Conflict Resolution* 53 (4): 570–597.

Fukumoto, K. 2015. "What Happens Depends on When It Happens: Copula-Based Ordered Event History Analysis of Civil War Duration and Outcome." *Journal of the American Statistical Association* 110 (509): 83–92.

Fukumoto, K. 2023a. Not Linearly Correlated, But Dependent: A Family of Normal Mode Copulas. arXiv:arXiv:2308.03541 [stat.ME].

Fukumoto, K. 2023b. Replication Data for: Normal Mode Copulas for Nonmonotonic Dependence. https://doi.org/10.7910/DVN/X94ITA, Harvard Dataverse, V1.

Genz, A., et al. 2021. "mvtnorm: Multivariate Normal and t Distributions." R Package Version 1.1–3. http://CRAN.R-project.org/package=mvtnorm

Gomes, M., R. Radice, J. C. Brenes, and G. Marra. 2019. "Copula Selection Models for Non-Gaussian Outcomes That Are Missing Not at Random." *Statistics in Medicine* 38 (3): 480–496.

Hankin, R. K. S. 2006. "Special Functions in R: Introducing the gsl Package." *R News* 6 (4).

Hofert, M., I. Kojadinovic, M. Maechler, and J. Yan. 2018. *Elements of Copula Modeling with R*. Gewerbestrasse: Springer.

Kocak, K., and Ö. Kıbrıs. 2023. "Social Media and Press Freedom." *British Journal of Political Science* 53 (1): 140–162.

Li, D. X. 2000. "On Default Correlation." *Journal of Fixed Income* 9 (4): 43–54.

Martin, L. W., and R. T. Stevenson. 2010. "The Conditional Impact of Incumbency on Government Formation." *American Political Science Review* 104 (3): 503–518.

Mason, T. D., J. P. WeingartenJr., and P. J. Fett. 1999. "Win, Lose, or Draw: Predicting the Outcome of Civil Wars." *Political Research Quarterly* 52 (2): 239–268.

Nelsen, R. B. 2006. *An Introduction to Copulas*, 2nd. ed. New York: Springer.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing. Version 4.1.3*. Vienna: R Foundation for Statistical Computing.

Sartori, A. E. 2003. "An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions." *Political Analysis* 11 (2): 111–138.

Schafer, J., R. Opgen-Rhein, V. Zuber, A. Miika Ahdesmaki, P. D. Silva, and K. Strimmer. 2021. "Corpcor: Efficient Estimation of Covariance and (Partial) Correlation." R Package Version 1.6.10. https://CRAN.R-project.org/package=corpcor

Trivedi, P. K., and D. M. Zimmer. 2007. *Copula Modeling: An Introduction for Practioners*. Boston: Publishers Inc.

Cite this article: Fukumoto, K. (2024). Normal Mode Copulas for Nonmonotonic Dependence. *Political Analysis*, 417–430. https://doi.org/10.1017/pan.2023.45