Technical and Allocative Inefficiency in Production Systems: A Vine Copula Approach*

Jian Zhai† **Artem Prokhorov**‡ **Robert James**§

July 2021

Abstract

Production systems that account for technical and allocative inefficiencies offer a natural way to model dependence using vine copulas. We construct such vine copulas using a recently proposed family of bivariate copulas that permit dependence between the magnitude (but not the sign) of allocative inefficiency and the magnitude of technical inefficiency. We show how to estimate such models and argue that they better reflect dependencies that arise in practice.

JEL Codes: C13
Key Words: vine copulas, allocative inefficiency, technical inefficiency

*Helpful comments from Christine Amsler, Brendan Beare, Peter Schmidt and the participants of SETA2019 and INFORMS2019 are gratefully acknowledged. Research for this paper was supported by a grant from the Russian Science Foundation (Project No. 20-18-00365). The use of the University of Sydneys high performance computing cluster, Artemis, is acknowledged.

†University of Sydney, Business School, NSW, Australia; email: jian.zhai@sydney.edu.au

‡University of Sydney, Australia, and CEBA, St.Petersburg State University, Russia, and CIREQ, University of Montreal, Canada; email: artem.prokhorov@sydney.edu.au

§University of Sydney, Business School, NSW, Australia; email: r.james@sydney.edu.au

1 Introduction

In this paper, we consider the following production system

$$\ln(y) = \alpha_0 + \sum_{j=1}^{J} \alpha_j \ln(x_j) + v - u, \tag{1}$$

$$\ln(x_1) - \ln(x_j) = \ln(\alpha_1 p_j / \alpha_j p_1) + \omega_j, \quad j = 2, ..., J, \tag{2}$$

where the first equation represents a Cobb-Douglas type production function of a generic production unit, with output y and inputs $x_j, j = 1, \ldots J$, and the set of the $J-1$ equations that follow the production function derive from the first-order conditions of cost minimization. They define the $J-1$ optimal input ratios in terms of input prices p_j. This is the system proposed by Schmidt and Lovell (1979, 1980) and most recently considered by Amsler et al. (2020).

The error terms in each of the equation have an economic interpretation. The two errors in first equation are a symmetric component v, which accounts for random factors affecting the production potential, e.g., weather, and an asymmetric component u, representing the technical inefficiency of production, that is the percentage by which the production unit falls short of the stochastic production frontier. The ω's represent allocative inefficiencies of the production unit. That is, they are the percentages by which the ratios x_1/x_j deviate from their cost-minimizing values. Clearly, production units may be technically inefficient due to allocative inefficiencies and vice versa, so u and ω's need to be allowed to depend on each other.

The choice of a parameterization for the dependence between u and ω's is not trivial. A naive approach would simply allow for non-zero correlation between u and ω's. This would mean that *either* too high *or* too low values of x_1/x_j, but not both at the same time, are associated with greater inefficiency. In fact, both negative and positive deviations from the cost-minimizing input ratios lead to higher u equally likely so what we need is a dependence parameterization that permits correlation between u and $|\omega|$, not between u and ω.

Schmidt and Lovell (1980) achieve this by using joint normality. They let $u = |u^*|$, where $u^* \sim N(0, \sigma_u^2)$ and assume that the errors $(u^*, \omega_2, \ldots, \omega_J)$ are jointly normal. They show that in this case u is uncorrelated with ω's but, by a result due to Nabeya (1951), u and $|\omega_j|$ have a non-zero correlation of the form

$$\frac{2\sigma_u \sigma_j}{\pi} \left[\sqrt{1-\rho^2} + \rho \arcsin(\rho) - 1 \right],$$

where σ_j is the standard deviation of ω_j and ρ is the correlation between u^* and ω_j.

Amsler et al. (2020) asked a more general question. They seek to characterize a family of copulas for which, given any set of marginals $F_u, F_{w_1}, \ldots, F_{w_J}$, the random variables u and $\omega_j, j = 1, \ldots, J$, are not correlated but the random variables u and $|\omega_j|$ are. They propose such a family based on the class of the Sarmanov copulas. They also derive the copula implied by the Schmidt and Lovell (1980) construction.

This paper builds on these results by introducing vine copulas into the construction of the joint distribution of $(u, \omega_1, \ldots, \omega_J)$. The vine copulas simplify the task of constructing multivariate joint densities with a desired dependence pattern by representing high-dimensional copula densities as products of bivariate conditional copulas. This is a natural representation in our setting because we are concerned with modelling bivariate dependences between u and ω_j. The joint distribution of $\omega_j, j = 2, \ldots, J$, will be permitted to remain conventional.

The proposed vine decomposition is relevant more generally in stochastic frontier models with endogeneity, where we wish to keep the marginal distribution of the inefficiency term pre-specified and allow dependence between it and the reduced form errors or endogenous regressors (see, e.g., Amsler et al., 2016; Tran and Tsionas, 2015). It is also of general statistical merit as a way of generating dependent random variables with a specific dependence structure, from uncorrelated random variables.

2 Vine Copulas

Copulas are multivariate distributions with uniform marginals (see, e.g., Nelsen, 2006, for an early introduction). The well-cited Sklar (1959) theorem states that any multivariate joint distribution of continuous random variables can be written uniquely as a copula function taking as arguments the univariate marginal distributions:[1]

$$H(z_1, \ldots, z_T) = C(F_1(z_1), ..., F_T(z_T)).$$

In terms of densities, we have

$$h(z_1, ..., z_T) = c(F_1(z_1), ..., F_T(z_T)) \prod_{t=1}^{T} f_t(z_t),$$

[1] Foundations of copulas were laid almost two decades earlier by Hoeffding (1940).

where $f_t(z_t)$ is the marginal density, corresponding to the cdf F_t, $c(\cdot, \ldots, \cdot)$ is the T-copula density corresponding to the copula function C, and $h(z_1, ..., z_T)$ is the joint density function corresponding to the joint cdf H.

A vine copula decomposition, proposed by Joe (1996), makes use of an equivalent representation of $h(z_1, ..., z_T)$ in terms of conditional densities (we provide details in Appendix A). For example, if $T = 4$, a canonical vine decomposition is

$$\begin{aligned} h(z_1, ..., z_4) &= c_{12}(F_1(z_1), F_2(z_2)) \cdot c_{13}(F_1(z_1), F_3(z_3)) \cdot c_{14}(F_1(z_1), F_4(z_4)) \\ &\times c_{23|1}(F(z_2|z_1), F(z_3|z_1)) \cdot c_{24|1}(F(z_2|z_1), F(z_4|z_1)) \\ &\times c_{34|12}(F(z_3|z_1, z_2), F(z_4|z_1, z_2)) \\ &\times \prod_{t=1}^{4} f_t(z_t), \end{aligned} \tag{3}$$

where a 4-copula density is decomposed into a product of six 2-copula densities, three of which are acting on unconditional marginal cdf's, two are acting on conditional cdf's with one variable in the conditioning set and one is acting on conditional cdf's with two variables in the conditioning set. The decomposition is particularly useful in settings with high T as a means of dimension reduction.

An important element of a vine construction is the conditional univariate cdf's used as arguments of the 2-copula densities. It is easy to see that for a single variable in the conditioning set, the conditional cdf can be written in terms of the corresponding copula as follows

$$F(z_i|z_j) = \frac{\partial C_{ij}(F_i(z_i), F_j(z_j))}{\partial F_j(z_j)}$$

For the case with more than one variable in the conditioning set, Joe (1996) shows that the following formula applies:

$$F(z_i|z_j, z_k, z_l) = \frac{\partial C_{ij|kl}(F(z_i|z_k, z_l), F(z_j|z_k, z_l))}{\partial F(z_j|z_k, z_l)}.$$

So it is important that the 2-copula densities can be integrated efficiently in each dimension.

The vine copula construction is known to be sufficiently general to capture a wide range of dependence even in extremely high dimensions (see, e.g., Spanhel and Kurz, 2019).

3 The APS copula family

Amsler et al. (2020) consider copulas of the form

$$c(\xi_1, \xi_2) = 1 + \theta g(\xi_1) h(\xi_2),$$

where $\int_0^1 g(s)ds = \int_0^1 h(s)ds = 0$ and where θ satisfies the restrictions that are necessary for c to be a density (see, e.g., Sarmanov, 1966). They characterize the functions $g(\xi_1)$ and $h(\xi_2)$ such that $cov(\xi_1, \xi_2) = 0$ while $cov(\xi_1, q(\xi_2)) \neq 0$ for some function $q(\cdot)$. They call a 2-copula with this property the APS-2 copula.

It turns out an APS-2 copula is obtained when $g(\xi_1) = 1 - 2\xi_1$ and $h(\xi_2) = 1 - k_q^{-1} q(\xi_2)$, where $q(\cdot)$ is integrable on $[0, 1]$, symmetric around $\xi = 1/2$, monotonically decreasing on $[0, \frac{1}{2}]$ and monotonically increasing on $[\frac{1}{2}, 1]$ and where $k_q = \int_0^1 q(s)ds$. For an APS-2 copula, they show that

$$cov(\xi_1, q(\xi_2)) = \frac{1}{6}\theta k_q^{-1} Var(q(\xi_2)),$$

so that θ is proportional to the correlation betweeen ξ_1 and $q(\xi_2)$. Amsler et al. (2020) show that if the original random variables $z_j = F^{-1}(\xi_j)$ have symmetric marginals with finite variance and are linked by an APS-2 copula, then $cov(z_1, z_2) = 0$ while generally $cov(z_1, |z_2|) \neq 0$.

The two members of the APS-2 family they consider are

$$\text{APS-2A:} \quad c(\xi_1, \xi_2) = 1 + \theta\left(1 - 2\xi_1\right)\left[1 - 12\left(\xi_2 - \tfrac{1}{2}\right)^2\right],$$
$$\text{APS-2B:} \quad c(\xi_1, \xi_2) = 1 + \theta\left(1 - 2\xi_1\right)\left[1 - 4\left|\xi_2 - \tfrac{1}{2}\right|\right],$$

for which they show that

$$\text{APS-2A:} \quad corr\left(\xi_1, \left(\xi_2 - \tfrac{1}{2}\right)^2\right) = \tfrac{2}{\sqrt{15}}\theta,$$
$$\text{APS-2B:} \quad corr\left(\xi_1, \left|\xi_2 - \tfrac{1}{2}\right|\right) = \tfrac{1}{3}\theta.$$

Extensions to $T > 2$ using 2-copulas are usually hard to achieve, due to numerous non-compatibility results showing that 2-copulas that act on 2-copulas generally do not produce 3- or 4-copulas (see, e.g., Nelsen, 2006, pp. 105-107). However, Amsler et al. (2020) show a remarkable result that a valid T-copula can be constructed using C_T^2 2-copulas as follows

$$c(\xi_1, \dots, \xi_T) = 1 + \sum_{1 \le i < j \le T} (c_{ij} - 1), \tag{4}$$

where c_{ij} is the 2-copula of ξ_i and ξ_j. For example, a valid 4-copula has the form

$$c(\xi_1, \dots, \xi_4) = 1 + (c_{12} - 1) + (c_{13} - 1) + (c_{14} - 1) + (c_{23} - 1) + (c_{24} - 1) + (c_{34} - 1).$$

An interesting property of such copulas is that the 2-copulas implied by them are the 2-copulas used to construct them. For example, for the 4-copula the implied 2-copulas are $c_{12}, c_{13}, c_{14}, c_{23}, c_{24}$ and c_{34}.

Based on this result, Amsler et al. (2020) define the APS-T copula as follows:

$$\begin{aligned}
\text{APS-}T: \qquad & c(\xi_1, \dots, \xi_T) = 1 + \textstyle\sum_{1 \le i < j \le T} (c_{ij} - 1), \\
\text{where} \qquad & c_{1j}(\xi_1, \xi_j) = \text{APS-2 copula density}, j = 2, \dots, T, \\
\text{and} \qquad & c_{kl}(\xi_k, \xi_l) = \text{bivariate Gaussian copula density}, k, l, \ne 1
\end{aligned}$$

Two members of this family would arise if we use APS-2A and APS-2B. The key feature of this family is that $cov(\xi_1, \xi_j) = 0$ while in general $cov(\xi_1, q(\xi_j)) \ne 0$, $j = 2, \dots, T$.

Two important limitations of the APS-T copula restrict the range of dependence between T random variables it can accommodate. First, the Sarmanov specification is a perturbation of independence which is generally not a comprehensive copula. For example, the Eyraud-Farlie-Gumbel-Morgenstern copula, which is a member of the Sarmanov class, cannot accommodate dependence outside the Kendall τ range of $[-2/9, 2/9]$.

Second, by construction, the APS-T family ($T > 2$) can accommodate no dependence beyond that implied by pairwise copulas. This follows from (4) since all lower-dimensional marginals of APS-T are also expressed in terms of c_{ij}. For example, c_{123} contains terms of the form $g(\xi_i)h(\xi_j), i, j = 1, 2, 3$, but no terms of the form $g(\xi_1)h(\xi_2)m(\xi_3)$, for some function m. This places additional restrictions on the type and strength of dependence that can be estimated using APS-T.

4 Vine copulas for the production system

We propose using vine copulas to construct the joint density of the error terms in system (1)-(2). We follow the convention and assume that v and ω_j are marginally normal and we assume that u is half-normal but any other asymmetric distribution would work. We use F and f to denote a cdf and pdf, respectively, and the subscripts will denote which error's distribution is in question, e.g., F_u, F_v are the cdfs of u and v, respectively, while F_j is the cdf of $\omega_j, j = 2, \ldots, J$.

For example, if $J = 4$, the vine decomposition (3) of the joint density of $(u, \omega_2, \omega_3, \omega_4)$ can be written as follows:

$$\begin{aligned} h(u, \omega_2, \omega_3, \omega_4) &= c_{12}(F_u(u), F_2(\omega_2)) \cdot c_{13}(F_u(u), F_3(\omega_3)) \cdot c_{14}(F_u(u), F_4(\omega_4)) \\ &\times c_{23|1}(F(\omega_2|u), F(\omega_3|u)) \cdot c_{24|1}(F(\omega_2|u), F(\omega_4|u)) \\ &\times c_{34|12}(F(\omega_3|u, \omega_2), F(\omega_4|u, \omega_2)) \\ &\times \prod_{j=1}^{3} f_j(\omega_j) f_u(u). \end{aligned} \tag{5}$$

We wish to preserve the key property of the APS copula family for the pairs $(u, \omega_j), j = 2, 3, 4$, so a natural choice of $c_{1j}, j = 2, 3, 4$, is an APS-2 copula. For the symmetric errors, we assume joint normality and hence a conventional choice for $c_{kl|1}$ and $c_{34|12}$ is the Gaussian copula.

A conceptual difference from the APS-4 copula is that the vine copula approach places virtually no restriction on the type of dependence (aside from that between u and ω_j) that can be accommodated. For example, no restriction is being placed on how c_{123} is or is not expressed in terms of c_{ij} only. Similarly, the vine approach is more general than the approach based on joint normality used by Schmidt and Lovell (1980).

The use of APS-2 provides a number of important computational advantages because the conditional distributions have simple closed-form expressions. In order to see this, differentiate the APS-2 copula function (or integrate the APS-2 copula density) with respect to the second argument. For example, for u and ω_2, the two members of the APS-2 family imply the following distributions

$$\begin{aligned} \text{APS-2A:} \quad C_{12}(F_u(u), F_2(\omega_2)) &= F_u(u) \cdot F_2(\omega_2) + \theta_{12} \cdot F_u(u) \cdot (1 - F_u(u)) \cdot F_2(\omega_2) \\ &\times [1 - (4F_2(\omega_2)^2 - 6F_2(\omega_2) + 3)], \end{aligned}$$

$$\text{APS-2B:} \quad C_{12}(F_u(u), F_2(\omega_2)) = \begin{cases} F_u(u) \cdot F_2(\omega_2) + \theta_{12} \cdot F_u(u) \cdot (1 - F_u(u)) \\ \times F_2(\omega_2) \cdot (2F_2(\omega_2) - 1), \qquad F_2(\omega_2) \leq \frac{1}{2}, \\ F_u(u) \cdot F_2(\omega_2) + \theta_{12} \cdot F_u(u) \cdot (1 - F_u(u)) \\ \times (F_2(\omega_2) - 1) \cdot (1 - 2F_2(\omega_2)), \quad F_2(\omega_2) > \frac{1}{2}. \end{cases}$$

Then, the required conditional distributions can be written as follows

$$\text{APS-2A:} \quad F(\omega_2|u) = F_2(\omega_2) + \theta_{12}(1 - 2F_u(u))(-4F_2(\omega_2)^3 + 6F_2(\omega_2)^2 + 2F_2(\omega_2))$$

and

$$\text{APS-2B:} \quad F(\omega_2|u) = \begin{cases} F_2(\omega_2) + \theta_{12}F_2(\omega_2)(2F_2(\omega_2) - 1)(1 - 2F_u(u), \quad F_2(\omega_2) \leq \frac{1}{2}, \\ F_2(\omega_2) + \theta_{12}(F_2(\omega_2) - 1)(1 - 2F_2(\omega_2))(1 - 2F_u(u)), \quad F_2(\omega_2) > \frac{1}{2} \end{cases}$$

The joint (conditional) normality assumption on ω's leads to a simple formula for conditional cdf's with two variables in the conditioning set. For example, the conditional copula of ω_2 and ω_3 given u can be written as follows

$$C_{23|1}\{F(\omega_2|u), F(\omega_3|u)\} = \Phi_2(\Phi^{-1}(F(\omega_2|u)), \Phi^{-1}(F(\omega_3|u)); \rho_{23}),$$

where Φ_2 denotes a bivariate normal cdf, Φ^{-1} denotes the inverse of a standard normal cdf and ρ_{23} is the parameter of a bivariate normal cdf (i.e., it is the correlation between $\Phi^{-1}(F(\omega_2|u))$ and $\Phi^{-1}(F(\omega_3|u))$). Thus, the conditional distribution function for ω_3 given u and ω_2 is

$$\begin{aligned} F(\omega_3|u, \omega_2) &= \frac{\partial C_{23|1}\{F(\omega_2|u), F(\omega_3|u)\}}{\partial F(\omega_2|u)} \\ &= \Phi\left(\frac{\Phi^{-1}(F(\omega_3|u)) - \rho\Phi^{-1}(F(\omega_2|u))}{\sqrt{1 - \rho_{23}^2}}\right). \end{aligned}$$

Other conditional distributions that serve as arguments of the bivariate copula densities in vine decompositions of the form (5) can be derived similarly.

5 Estimation of parameters and technical inefficiencies

Eventually, we construct the joint density of u, v and ω's and use it to form a likelihood and to find the parameter vector that maximizes it. It is customary to assume that v, being the random noise, is independent of u and w's and normal. Then, the joint density of υ, u, ω's can be written as follows

$$h(\upsilon, u, \omega_2, \ldots, \omega_J) = f(\upsilon) \cdot c(F(u), F_2(\omega_2), \ldots, F_J(\omega_J)) \cdot f(u) \cdot \prod_{j=2}^{J} f_j(\omega_j),$$

where the copula term assumes a vine decomposition similar to (5).

As noted by Amsler et al. (2020), u and v are not observed separately, and in order to apply MLE we need the joint density of $v - u$ and ω's. Let $\varepsilon = v - u$. Then, as long as $h(\upsilon, u, \omega_2, \ldots, \omega_J)$ is available, Amsler et al. (2020) obtain the required joint density using expectation over the distribution of u as follows

$$\begin{aligned} f(\varepsilon, \omega_2, \ldots, \omega_J) &= \int h(u + \varepsilon, u, \omega_2, \ldots, \omega_J) du \\ &= \prod_{j=2}^{J} f_j(\omega_j) \cdot E_u[c(F(u), F_2(\omega_2), \ldots, F_J(\omega_J)) \cdot f_v(u + \varepsilon)], \end{aligned} \tag{6}$$

where $f(\varepsilon, \omega_2, \ldots, \omega_J)$ is the joint density of $v - u$ and ω's and E_u is expectation with respect to $f(u)$. The joint density can be approximated to a desired precision by taking an average of a large number of random draws of u.

The vine decomposition permits an additional representation of $f(\varepsilon, \omega_2, \ldots, \omega_J)$ given by (5) and the assumption that c_{1j}'s are APS-2 and the conditional copulas linking ω's are normal. For example, for $J = 3$,

$$\begin{aligned} f(\varepsilon, \omega_2, \omega_3) = f(\omega_2) \cdot f(\omega_3) \cdot \big\{ & E_u[c_{23|1} \cdot f_v(u + \varepsilon)] \\ & + [\theta_{12} \cdot h(F(\omega_2)) + \theta_{13} \cdot h(F(\omega_3))] \times E_u[g(F(u)) \cdot c_{23|1} \cdot f_v(u + \varepsilon)] \\ & + \theta_{12} \cdot \theta_{13} \cdot h(F(\omega_2)) \cdot h(F(\omega_3)) \times E_u[g^2(F(u)) \cdot c_{23|1} \cdot f_v(u + \varepsilon)] \big\}, \end{aligned}$$

where $c_{23|1}$ denotes the Gaussian copula density evaluated at $\xi_j = F(\omega_j | u), j = 2, 3$. The Gaussian density can be written as follows

$$c_{23|1}(\xi_2, \xi_3) = \frac{1}{\sqrt{1 - \rho_{23}^2}} \exp\left(-\frac{\rho_{23}^2 (c_2^2 + c_3^2) - 2\rho_{23} \cdot c_2 \cdot c_3}{2(1 - \rho_{23}^2)} \right),$$

where $c_j = \Phi^{-1}(\xi_j)$. As before, we can use $g(\xi_1) = 1 - 2\xi_1$ and $h(\xi_j) = 1 - 12(\xi_j - \frac{1}{2})^2$ for APS-2A and $h(\xi_j) = 1 - 4|\xi_j - \frac{1}{2}|$ for APS-2B, $j = 2, 3$.

A Maximum Simulated Likelihood Estimation (MSLE) is based on maximizing the following log-likelihood function

$$\ln L(\beta) = \sum_{i=1}^{n} \ln f(\varepsilon_i, \omega_{i2}, \dots, \omega_{iJ}),$$

where β contains all the model parameters to be estimated and where $\varepsilon_i = \ln y_i - \alpha_0 - \sum_{j=1}^{J} \alpha_j \ln x_{ij}$ and $\omega_{ij} = \ln(x_{i1}) - \ln(x_{ij}) - \ln(\alpha_1 p_{ij}/\alpha_j p_{i1}), j = 2, \dots, J$. The parameter vector β contains the parameters of the production function as well as the distributional parameters of the error terms, including the copula parameters. For example, if $J = 3$, this includes four parameters in the production function $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$, six parameters from the marginal distributions of the error terms (means μ_j and variances σ_j^2 of each $\omega_j, j = 2, 3$, variance σ_v^2 of v and variance σ_u^2 of u^*), plus three dependence parameters (θ_{12} and θ_{13} are the parameters of the APS-2 copulas for (u, ω_2) and (u, ω_3), respectively, and ρ_{23} is the Gaussian copula parameters for (ω_2, ω_3)). Note that if $\theta_{1j} \neq 0$ then $corr[F(u), q(F_j(\omega_j))] \neq 0$ even if $corr(u, \omega_j) = 0$.

Once β is estimated, we can obtain inefficiency scores $\hat{u}_i$ using the well known formula due to Jondrow et al. (1982):

$$\hat{u}_i = E(u|\varepsilon_i) = \sigma_* \left[\frac{\phi(b_i)}{1 - \Phi(b_i)} - b_i \right],$$

where $\sigma_* = \sigma \frac{\lambda}{1+\lambda^2}, b_i = \varepsilon_i \lambda / \sigma$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\lambda = \sigma_u / \sigma_v$. We use the residuals $\hat{\varepsilon}_i = \ln y_i - \hat{\alpha}_0 - \sum_{j=1}^{J} \hat{\alpha}_j \ln x_{ij}$ in place of ε_i, to evaluate the conditional expectation. However, as noted by Amsler et al. (2016), the availability of the residuals $\hat{\omega}_j$ allows for a better, more precise, prediction of u. Unfortunately, their approach does not work here, as they allow for dependence between ω and v, not between ω and u.

In our settings it *is* possible to use the approach proposed by Amsler et al. (2014, Section 5.2) for panel stochastic frontiers. They make use of the fact that, once we estimate the model, we know the joint distribution of $(u, \varepsilon, \omega_2, \dots, \omega_J)$ and can simulate from it. Then, we can use any nonparametric smoother to estimate the conditional expectation. Let $J = 3$ and let $s = 1, \dots, S$ index the draws from the joint distribution. Then, the Nadaraya-Watson estimator can be written as follows

$$\tilde{u}_i = E(u|\varepsilon_i, \omega_{i2}, \omega_{i3}) = \frac{\sum_{s=1}^{S} u_s K(\frac{\varepsilon_s - \varepsilon_i}{h}) K(\frac{\omega_{s2} - \omega_{i2}}{h}) K(\frac{\omega_{s3} - \omega_{i3}}{h})}{\sum_{s=1}^{S} K(\frac{\varepsilon_s - \varepsilon_i}{h}) K(\frac{\omega_{s2} - \omega_{i2}}{h}) K(\frac{\omega_{s3} - \omega_{i3}}{h})},$$

where $K(\cdot)$ is a kernel function and h is a bandwidth parameter. In practice, one would often use the Gaussian kernel with the rule-of-thumb value of $h = \frac{4\hat{\sigma}^5}{3n} = 1.06\hat{\sigma}n^{-1/5}$, where $\hat{\sigma}$ is the standard deviation of the relevant draws. As before, we would evaluate this estimator at the values of the residuals but in our settings, we have both the residuals $\hat{\varepsilon}_i$, defined earlier, and $\hat{\omega}_{ij} = \ln(x_{i1}) - \ln(x_{ij}) - \ln(\hat{\alpha}_1 p_{ij}/\hat{\alpha}_j p_{i1})$. It is a standard result in nonparametrics that, under $S \to \infty$, $h \to 0$ and $Sh^{J+1} \to \infty$, the new estimator $\tilde{u}_i$ converges to $E(u|\varepsilon_i, \omega_{i2}, \omega_{i3})$.

6 Monte Carlo Simulations

We evaluate the performance of the proposed vine construction in terms of MSE. The production function and endogenous regressor equations are

$$y = \alpha + x_1\beta_1 + x_2\beta_2 + \upsilon - u$$
$$x_j = z_j\gamma + \omega_j, \quad j = 1, 2,$$

where z_1 and z_2 are generated independently as χ^2_2. We generate error terms as follows: $\omega_j \sim N(0, \sigma_j^2)$, $j = 1, 2$, jointly normal with correlation ρ; $u \sim |N(0, \sigma_u^2)|$ and $\upsilon \sim N(0, \sigma_\upsilon^2)$ that is independent of u and ω_j.[2] The copula used to generate dependence between u and $|\omega_j|$ is APS-2B with parameters θ_1 for (u, ω_1) and θ_2 for (u, ω_2).

First, we generate three dependent uniforms from the canonical vine copula using the algorithm described, e.g., by Aas et al. (2009) and Czado (2019). Then, the respective inverse distribution functions are used to generate the vector (u, ω_1, ω_2).

The true parameter values are $\alpha = \boldsymbol{\beta} = 0.5$, $\sigma_\upsilon^2 = \sigma_u^2 = \sigma_{\omega_j}^2 = 1$, $\gamma = 1$ and $\rho = 0.4$. We consider three combinations of copula parameter values $(\theta_1, \theta_2) \in \{(0.3, 0.1), (0.45, 0.45), (0.8, 0.7)\}$, corresponding to low, medium and high dependence between u and $|\omega_j|$, respectively, and two sample sizes $n \in \{500, 1000\}$. We conduct 1000 replications. A sample size of 100 is used to evaluate the expectation in (6)

We compare how the vine-copula based estimators perform in terms of parameter-specific and aggregate MSE with three other estimators of the parameters in the marginals and in the production equation. Table 1 reports the results. Vine2A and Vine2B are the proposed estimators which use the APS-2A and APS-2B copulas as described in the previous section. APS3A and APS3B are the estimators that use the high-dimensional APS copulas rather than vines. QMLE is the estimator based on the assumption of independence between u and ω's (Schmidt and Lovell, 1979). Gaussian is the estimator that assumes the Gaussian copula for (u, ω_1, ω_2).

[2]As an alternative design we have used the setup of Tran and Tsionas (2015) and obtained qualitatively similar results. They are available upon request.

Table 1: MSE comparisons

	(i) n = 500						(ii) n = 1000					
	QMLE	Gaussian	APS3A	APS3B	Vine2A	Vine2B	QMLE	Gaussian	APS3A	APS3B	Vine2A	Vine2B
	$\theta_1 = 0.3, \theta_2 = 0.1$						$\theta_1 = 0.3, \theta_2 = 0.1$					
α	0.3394	0.4889	0.3171	0.3563	0.3670	0.3969	0.2342	0.4895	0.3514	0.3767	0.3729	0.3908
β_1	0.0236	0.0249	0.0240	0.0236	0.0237	0.0235	0.0167	0.0178	0.0168	0.0167	0.0167	0.0168
β_2	0.0238	0.0255	0.0245	0.0244	0.0242	0.0241	0.0162	0.0176	0.0164	0.0163	0.0163	0.0162
σ_u^2	0.2040	0.2531	0.1997	0.2180	0.2211	0.2270	0.1521	0.2506	0.2109	0.2232	0.2194	0.2225
σ_v^2	0.5653	0.7032	0.5272	0.5862	0.5956	0.6168	0.4204	0.6883	0.5712	0.6071	0.5982	0.6058
σ_1^2	0.1235	0.0648	0.0654	0.0649	0.0650	0.0646	0.0846	0.0446	0.0446	0.0447	0.0447	0.0446
σ_2^2	0.1136	0.0718	0.0615	0.0603	0.0610	0.0606	0.0828	0.0513	0.0443	0.0441	0.0442	0.0441
γ	0.0115	0.0115	0.0117	0.0116	0.0116	0.0116	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080
Total	1.4047	1.6437	1.2310	1.3453	1.3692	1.4252	1.0150	1.5676	1.2637	1.3369	1.3204	1.3488
$E[u\|\varepsilon]$	0.6302	0.7349	0.6152	0.6416	0.6480	0.6679	0.5731	0.7372	0.6410	0.6588	0.6560	0.6679
$E[u\|\varepsilon, \omega_1, \omega_2]$		0.7363	0.6177	0.6439	0.6506	0.6700		0.7402	0.6450	0.6628	0.6608	0.6717
	$\theta_1 = 0.45, \theta_2 = 0.45$						$\theta_1 = 0.45, \theta_2 = 0.45$					
α	0.3410	0.4936	0.2306	0.2757	0.2721	0.2798	0.2357	0.4745	0.2347	0.2670	0.2220	0.2184
β_1	0.0237	0.0254	0.0252	0.0242	0.0239	0.0238	0.0167	0.0181	0.0175	0.0168	0.0166	0.0166
β_2	0.0238	0.0255	0.0254	0.0245	0.0240	0.0239	0.0163	0.0176	0.0172	0.0164	0.0163	0.0163
σ_u^2	0.2046	0.2581	0.1593	0.1821	0.1761	0.1830	0.1523	0.2447	0.1579	0.1745	0.1505	0.1488
σ_v^2	0.5679	0.7186	0.3916	0.4830	0.4694	0.4817	0.4218	0.6744	0.4093	0.4700	0.4070	0.3995
σ_1^2	0.1235	0.0648	0.0703	0.0648	0.0654	0.0647	0.0846	0.0446	0.0484	0.0447	0.0447	0.0447
σ_2^2	0.1136	0.0713	0.0652	0.0609	0.0611	0.0608	0.0828	0.0513	0.0478	0.0443	0.0443	0.0440
γ	0.0115	0.0115	0.0119	0.0116	0.0115	0.0114	0.0080	0.0080	0.0081	0.0081	0.0080	0.0080
Total	1.4096	1.6688	0.9795	1.1268	1.1035	1.1291	1.0180	1.5331	0.9408	1.0418	0.9095	0.8962
$E[u\|\varepsilon]$	0.6318	0.7377	0.5645	0.5914	0.5870	0.5928	0.5740	0.7279	0.5727	0.5918	0.5679	0.5664
$E[u\|\varepsilon, \omega_1, \omega_2]$		0.7393	0.5648	0.5910	0.5859	0.5918		0.7309	0.5744	0.5933	0.5691	0.5668
	$\theta_1 = 0.8, \theta_2 = 0.7$						$\theta_1 = 0.8, \theta_2 = 0.7$					
α	0.3423	0.4873	0.1586	0.1792	0.1456	0.1692	0.2390	0.4787	0.1365	0.1569	0.1063	0.1115
β_1	0.0238	0.0258	0.0265	0.0251	0.0238	0.0238	0.0167	0.0183	0.0189	0.0179	0.0168	0.0165
β_2	0.0238	0.0259	0.0263	0.0256	0.0242	0.0244	0.0163	0.0180	0.0187	0.0175	0.0166	0.0163
σ_u^2	0.2056	0.2578	0.1299	0.1340	0.1191	0.1354	0.1527	0.2516	0.1145	0.1203	0.0894	0.0966
σ_v^2	0.5701	0.7165	0.2488	0.3098	0.2756	0.3398	0.4234	0.6919	0.2343	0.2910	0.2172	0.2382
σ_1^2	0.1235	0.0648	0.0842	0.0668	0.0673	0.0645	0.0846	0.0446	0.0588	0.0461	0.0456	0.0447
σ_2^2	0.1136	0.0715	0.0784	0.0623	0.0628	0.0609	0.0828	0.0513	0.0595	0.0455	0.0445	0.0438
γ	0.0115	0.0115	0.0120	0.0118	0.0116	0.0114	0.0080	0.0080	0.0086	0.0082	0.0080	0.0079
Total	1.4141	1.6611	0.7647	0.8146	0.7300	0.8294	1.0234	1.5623	0.6499	0.7034	0.5443	0.5756
$E[u\|\varepsilon]$	0.6330	0.7348	0.5320	0.5422	0.5309	0.5409	0.5756	0.7319	0.5291	0.5381	0.5216	0.5241
$E[u\|\varepsilon, \omega_1, \omega_2]$		0.7366	0.5262	0.5341	0.5187	0.5269		0.7349	0.5245	0.5317	0.5092	0.5108

We can see that near independence (upper panel) all six estimators show similar performance for both sample sizes, with QMLE being not much worse and at times better than the other estimators in terms of aggregate MSE. As dependence increases (middle and lower panels) the two new estimators Vine2A and Vine2B behave similarly to APS3A and APS3B and dominate QMLE. In the case of strongest dependence (lower panel), Vine2A and Vine2B show superior performance even compared with APS3A and APS3B (particularly for the larger sample size), way ahead of QMLE and Gaussian. Parameter-specific MSEs show that this behavior is not limited to just a few parameters but is prevalent uniformly across all parameters. It is perhaps remarkable that the MSE is concentrated in the estimates of σ_v^2, regardless of the estimator, sample size, or dependence strength. The Gaussian estimator is dominated by all the other estimators.

The last two lines of each panel contain the MSE of the two variants of inefficiency predictions. We have the set of u_i's used in the data generating process in a given iteration. In addition, as a result of the estimation, on each iteration we also have a set of predictions $\hat{u}_i$ and $\tilde{u}_i$ using the two conditioning sets. For each iteration, we calculate the MSEs

$$\frac{1}{n}\sum_{i=1}^{n}(u_i - \hat{u}_i)^2, \qquad \frac{1}{n}\sum_{i=1}^{n}(u_i - \tilde{u}_i)^2.$$

The line showing $E(u|\varepsilon, \omega_2, \omega_3)$ is obtained using the average of the latter values over 1000 iterations. The line showing $E(u|\varepsilon)$ is obtained using the former. Perhaps, surprisingly, for all dependence strengths, we obtain very similar MSEs even though one estimator is nonparametric and the other uses an analytic expression.

7 Empirical Illustration

We illustrate the use of our estimator using electricity generation data from 111 privately-owned steam-electric power plants constructed in the US between 1947 and 1965 (see Cowing, 1970, 1974, for details). The output is measured in 10^6 KWh of electricity generated in the first year of operation, inputs are capital as measured by actual cost of construction, fuel as measured in BTU of actual consumption of coal, oil and gas in the first year, and labor as measured by the total number of employees times 2000 hours. We also have input prices: firm's bond rate prior to plant construction, actual price of a BTU of fuel and regional industry salary rate averaged over two years prior to plant opening. Summary statistics for the data are given in Table 2, where output and inputs have been logged.

Table 2: Descriptive Statistics for Electricity Generation Data

	Mean	Median	St.D.	Min	Max
Output	6.834	6.915	0.991	3.638	8.703
Capital	16.859	16.919	0.775	14.542	18.374
Fuel	16.094	16.138	0.888	13.346	17.772
Labor	11.655	11.678	0.503	10.086	12.725
Price of Capital	-3.329	-3.387	0.192	-3.594	-2.947
Price of Fuel	-1.337	-1.241	0.313	-2.797	-0.877
Price of Labor	0.800	0.829	0.247	0.300	1.278

We use the Cobb-Douglas specification. As a benchmark, we reproduce the estimates of Schmidt and Lovell (1980) [SL80] which use joint normality. In addition, we provide estimates obtained under the assumption of Amsler et al. (2016) [APS16] which is that υ is correlated with ω, but u is independent of υ and ω. We also report the estimates based on Amsler et al. (2020) which use the APS-3 copulas [APS3A and APS3B]. Table 3 contains the results. The proposed estimators (Vine APS2A and Vine APS2B) are in the last four columns.

We start by noticing that the estimated input-output elasticities are similar for all the estimators that allow for dependence between u and ω's [SL80, APS3A&B, APS2A&B+Gaussian] and different from APS16. Moreover, there are substantial differences in the dependence parameter estimates. APS3A&B and Vine APS2A&B show positive dependence between u and ω_2 and negative between u and ω_3 while SL80 shows positive dependence for both. The correlations for all copula-based estimators are, for the most part, large in magnitude but statistically insignificant while SL80 shows a very weak but statistically significant correlation.

Finally, we note the similarity in the three versions of inefficiency estimates. The statistics $E[u|\varepsilon]$ and $V[u|\varepsilon]$ are the conditional mean and variance, computed as averages (over observations) of the closed form expressions in Jondrow et al. (1982) and Bera and Sharma (1999), respectively. The estimates $\tilde{E}[u|\varepsilon]$ and $\tilde{V}[u|\varepsilon]$ are the nonparametric (Nadaraya-Watson) versions of the same quantities. The estimates $\tilde{E}[u|\varepsilon, \omega_1, \omega_2]$ and $\tilde{V}[u|\varepsilon, \omega_1, \omega_2]$ use Nadaraya-Watson for the enlarged conditioning set. The SL80 and APS16 values are visibly different from the estimators based on the APS copula. Again, perhaps surprisingly, it makes little difference whether the estimator is parametric or nonparametric and whether we condition on ε or $(\varepsilon, \omega_2, \omega_3)$.

8 Conclusion

Production systems considered in this paper have an inherent structure of error terms. We proposed a vine copula construction that makes use of this structure and accounts for dependence between technical and allocative inefficiency. We argue that this construction is natural and achieves a more comprehensive coverage of dependence than the multivariate APS-copula proposed by Amsler et al. (2020). We also use this structure to implement a new way to estimate technical inefficiency scores, permitted by the fact that we can condition on more error terms.

9 Appendix: C-vine decomposition

See Bedford and Cooke (2001); Aas et al. (2009) for details of the general case. For $T = 4$, using conditioning, we can write any density as

$$h(z_1, ..., z_4) = f_1(z_1) \cdot f(z_2|z_1) \cdot f(z_3|z_1, z_2) \cdot f(z_4|z_1, z_2, z_3), \tag{7}$$

where the conditional densities $f(\cdot|\cdot)$ are represented in terms of copulas and marginals as follows:

$$\begin{aligned}
f(z_2|z_1) &= c_{12}\{F_1(z_1), F_2(z_2)\} \cdot f_2(z_2), \\
f(z_3|z_1, z_2) &= \frac{f(z_2, z_3|z_1)}{f(z_2|z_1)} \\
&= \frac{c_{23|1}(F(z_2|z_1), F(z_3|z_1)) \cdot f(z_3|z_1) \cdot f(z_2|z_1)}{f(z_2|z_1)} \\
&= c_{23|1}\{F(z_2|z_1), F(z_3|z_1)\} \cdot f(z_3|z_1) \\
&= c_{23|1}\{F(z_2|z_1), F(z_3|z_1)\} \cdot c_{13}\{F_1(z_1), F_3(z_3)\} \cdot f_3(z_3), \\
f(z_4|z_1, z_2, z_3) &= \frac{f(z_3, z_4|z_1, z_2)}{f(z_3|z_1, z_2)} \\
&= \frac{c_{34|12}\{F(z_3|z_1, z_2), F(z_4|z_1, z_2)\} \cdot f(z_3|z_1, z_2) \cdot f(z_4|z_1, z_2)}{f(z_3|z_1, z_2)} \\
&= c_{34|12}\{F(z_3|z_1, z_2), F(z_4|z_1, z_2)\} \cdot f(z_4|z_1, z_2) \\
&= c_{34|12}\{F(z_3|z_1, z_2), F(z_4|z_1, z_2)\} \cdot \frac{f(z_2, z_4|z_1)}{f(z_2|z_1)} \\
&= c_{34|12}\{F(z_3|z_1, z_2), F(z_4|z_1, z_2)\} \\
&\times \frac{c_{24|1}\{F(z_2|z_1), F(z_4|z_1)\} \cdot f(z_2|z_1) \cdot f(z_4|z_1)}{f(z_2|z_1)} \\
&= c_{34|12}\{F(z_3|z_1, z_2), F(z_4|z_1, z_2)\} \cdot c_{24|1}\{F(z_2|z_1), F(z_4|z_1)\} \cdot f(z_4|z_1) \\
&= c_{34|12}\{F(z_3|z_1, z_2), F(z_4|z_1, z_2)\} \cdot c_{24|1}\{F(z_2|z_1), F(z_4|z_1)\} \\
&\times c_{14}\{F_1(z_1), F_4(z_4)\} \cdot f_4(z_4).
\end{aligned}$$

Substituting into (7) produces the result in the main text.

There are clearly more than one way of applying the conditioning argument (see, e.g., Aas et al., 2009). The one we show corresponds to what is known as a C-vine (canonical vine) decomposition. A D-vine (drawable vine) decomposition, for example, has the form $h_{1234} = f_1 f_2 f_3 f_4 c_{12} c_{23} c_{34} c_{13|2} c_{24|3} c_{14|23}$. However, such alternative decompositions do not align with our goal of representing the dependence between u and each w_j via c_{1j}.

Table 3: MLE of Production Function Parameters

	SL80		APS16		APS3A		APS3B		Vine APS2A		Vine APS2B	
	Est	Std Err	Est	Std Err	Est	Std Err	Est	Std Err	Est	Std Err	Est	Std Err
α	-11.2700	0.2510	-11.6839	0.3865	-11.4126	0.2315	-11.4525	0.2272	-11.4482	0.2477	-11.4675	0.2542
β_1	0.0428	0.0246	0.2248	0.0142	0.0498	0.0227	0.0495	0.0234	0.0430	0.0250	0.0471	0.0004
β_2	1.0754	0.0272	0.8625	0.0251	1.0751	0.0232	1.0712	0.0230	1.0793	0.0247	1.0740	0.0197
β_3	0.0137	0.0319	0.0805	0.0060	0.0145	0.0271	0.0239	0.0174	0.0220	0.0286	0.0253	0.0301
μ_1	1.9861	0.6052			1.8481	0.4948	1.8271	0.5093	1.9926	0.6236	1.8920	0.0740
μ_2	-0.0526	2.4744			-0.1740	1.9530	0.3097	0.8434	0.3675	1.4985	0.4185	1.1890
σ_u^2	0.0119	0.0036	0.0089	0.0058	0.0083	0.0030	0.0080	0.0029	0.0086	0.0029	0.0086	0.0028
σ_v^2	0.0020	0.0009	0.0157	0.0048	0.0037	0.0010	0.0038	0.0011	0.0036	0.0011	0.0035	0.0010
σ_2^2	0.3366	0.0435	0.3503	0.0414	0.3437	0.0351	0.3416	0.0354	0.3456	0.0360	0.3391	0.0354
σ_3^2	0.5901	0.1010	0.5923	0.1007	0.5956	0.0980	0.5987	0.0992	0.5684	0.0958	0.5763	0.0937
ρ_{23}	0.2100	0.0577	0.2159	0.0549	0.5157	0.1024	0.5147	0.1009	0.4790	0.0998	0.4627	0.1043
σ_{2u}	0.0119	0.0036										
σ_{3u}	0.0148	0.0242										
σ_{2v}			0.0157	0.0048								
σ_{3v}			-0.0492	0.0125								
θ_{12}					0.5907	0.3489	0.7456	0.4774	0.3597	0.4176	0.0198	0.5956
θ_{13}					-0.4088	0.3688	-0.5465	0.5512	-0.3845	0.3751	-0.4526	0.5429
$E[u\|\varepsilon]$	0.0852		0.0820		0.0690		0.0688		0.0714		0.0729	
$V[u\|\varepsilon]$	0.0013		0.0029		0.0015		0.0015		0.0015		0.0015	
$\tilde{E}[u\|\varepsilon]$	0.0817		0.0738		0.0682		0.0679		0.0702		0.0713	
$\tilde{V}[u\|\varepsilon]$	0.0011		0.0027		0.0017		0.0017		0.0017		0.0018	
$\tilde{E}[u\|\varepsilon,\omega_2,\omega_3]$	0.0813		0.0207		0.0662		0.0664		0.0686		0.0711	
$\tilde{V}[u\|\varepsilon,\omega_2,\omega_3]$	0.0010		0.0013		0.0016		0.0016		0.0017		0.0017	

References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009): "Pair-copula constructions of multiple dependence," *Insurance: Mathematics and Economics*, 44, 182 – 198.

Amsler, C., A. Prokhorov, and P. Schmidt (2014): "Using copulas to model time dependence in stochastic frontier models," *Econometric Reviews*, 33, 497–522.

——— (2016): "Endogeneity in stochastic frontier models," *Journal of Econometrics*, 190, 280–288.

——— (2020): "A New Family of Copulas, with Application to Estimation of a Production Frontier System," *Journal of Productivity Analysis.*

Bedford, T. and R. M. Cooke (2001): "Probability density decomposition for conditionally dependent random variables modeled by vines," *Annals of Mathematics and Artificial intelligence*, 32, 245–268.

Bera, A. K. and S. C. Sharma (1999): "Estimating production uncertainty in stochastic frontier production function models," *Journal of Productivity Analysis*, 12, 187–210.

Cowing, T. (1970): *unpublished Ph.D. dissertation, University of California, Berkeley.*

Cowing, T. G. (1974): "Technical Change and Scale Economies in an Engineering Production Function: The Case of Steam Electric Power," *The Journal of Industrial Economics*, 23, 135–152.

Czado, C. (2019): "Analyzing dependent data with vine copulas," *Lecture Notes in Statistics, Springer.*

Hoeffding, W. (1940): "Masstabinvariante korrelationstheorie." *Schriften des Mathematischen Instituts und des Instituts fuer Angewandte Mathematik der Universitat Berlin*, 5, 179–233.

Joe, H. (1996): "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters," *in L. Rschendorf, B. Schweizer and M.D. Taylor (eds.), Distributions with Fixed Marginals and Related Topics, IMS Lecture Notes Monograph Series, Institute of Mathematical Statistics.*

Jondrow, J., C. K. Lovell, I. S. Materov, and P. Schmidt (1982): "On the estimation of technical inefficiency in the stochastic frontier production function model," *Journal of econometrics*, 19, 233–238.

Nabeya, S. (1951): "Absolute moments in two-dimensional normal distribution," *Annals of the Institute of Statistical Mathematics*, 3, 2–6.

Nelsen, R. B. (2006): "An introduction to copulas, 2nd," *New York: Springer Science Business Media*.

Sarmanov, O. (1966): "Generalized Normal Correlation and Two-Dimensional Frechet Classes," *Doklady (Soviet Mathematics)*, 168, 596–599.

Schmidt, P. and C. K. Lovell (1979): "Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers," *Journal of econometrics*, 9, 343–366.

——— (1980): "Estimating stochastic production and cost frontiers when technical and allocative inefficiency are correlated," *Journal of Econometrics*, 13, 83–100.

Sklar, M. (1959): "Fonctions de repartition an dimensions et leurs marges," *Publ. inst. statist. univ. Paris*, 8, 229–231.

Spanhel, F. and M. S. Kurz (2019): "Simplified vine copula models: Approximations based on the simplifying assumption," *Electronic Journal of Statistics*, 13, 1254–1291.

Tran, K. C. and E. G. Tsionas (2015): "Endogeneity in stochastic frontier models: Copula approach without external instruments," *Economics Letters*, 133, 85–88.