The Australian Journal of
Agricultural and Resource Economics
Journal of the Australian Agricultural and Resource Economics Society

Australian Journal of Agricultural and Resource Economics, 58, pp. 90–110

A copula-based approach to the simultaneous estimation of group and meta-frontiers by constrained maximum likelihood

Alexandre Repkine†

Existing approaches to the meta-frontier estimation consist of two stages where the estimates of the local frontier parameters obtained in the first step are used to estimate meta-frontier parameters by means of a linear or quadratic minimisation procedure in the second. Since it was shown by Schmidt (*Review of Economics and Statistics* 58: 238) that the second step is equivalent to constrained maximisation of a likelihood function, we extend this idea and offer a copula-based approach to the estimation of the parameters of both meta- and group frontiers in a one-step setting. In this way, we ensure a single data-generating mechanism for the estimated parameters, expand the set of potential meta-frontiers and account for the fact that shocks to the individual production units may be correlated with shocks to the local technological environment as a whole. We apply our estimation methodology to a data set on the world agriculture and find that the deviations from the group frontiers are positively correlated with deviations from the meta-frontier, which is a conclusion that is impossible to reach without accounting for stochastic dependence between the two deviation types represented by a copula.

Key words: constrained maximum likelihood, meta-frontiers, stochastic frontiers, technical efficiency.

1. Introduction

Stochastic frontier framework (SFA) introduced independently by Aigner *et al*. (1977) and Meeusen and van den Broeck (1977) allows one to measure and compare the extent of efficiency with which the production inputs are used. Battese and Coelli (1988) provide an SFA-based estimation procedure that uses data on production inputs and output levels in order to infer levels of productive efficiency. The comparison of decision-making units in terms of their productive efficiency may become biased if these units belong to different technological environments. O'Donnell *et al*. (2008) are recognising this issue by introducing the concept of a meta-frontier overarching the individual group frontiers. Distance between the meta-frontier and the group frontier can be thought of as a measure of the extent to which the local production environment is restricted due to, for example, lack of economic infrastructure, institutions and general development of the

†Alexandre Repkine (email: repkine@konkuk.ac.kr) is at Economics Department, College of Commerce and Economics, Konkuk University, 120 Neungdong-Ro, Gwangjin-Gu, Seoul, 143-701, South Korea.

doi: 10.1111/1467-8489.12026

economic environment. From the managerial perspective, it does not appear rational to channel scarce resources into the incentive schemes that would increase productive efficiency in a technologically restrictive environment.

The reason why meta-frontiers have so far been mostly estimated by the linear or quadratic programming techniques is that the stochastic frontier methodology (SF) applied to the pooled sample will not necessarily result in a meta-frontier that overarches the group frontiers at each observation point, as has been demonstrated by Battese and Rao (2002). For that reason, Battese *et al*. (2004) along with O'Donnell *et al*. (2008) suggest estimating the meta-frontiers by minimising the sum of the distances between meta- and group frontiers or these distances' squares subject to the constraint that the meta-frontier overarches group frontiers at all observed combinations of the input factors. The above minimisation is performed by applying the quadratic or linear programming (QLP) methodology.

Schmidt (1976) implies that, in fact, the QLP approach to the estimation of meta-frontiers is equivalent to the *constrained maximization of a likelihood function* corresponding to a meta-frontier model where distances between meta- and group frontiers are random variables distributed either exponentially or half-normally. In the former case, the constrained maximum likelihood (CML) estimation is equivalent to the minimisation of the absolute distances between meta- and group frontiers, while in the latter case, the CML procedure results in the meta-frontier parameter values obtained by minimising the sum of the squared distances.

The CML approach to the estimation of meta-frontiers is advantageous for several reasons. First, it is an estimation procedure that is based on a statistical model of the meta-frontier rather than a computational algorithm minimising the Euclidean distance between the meta- and group frontiers. Second, the CML approach accommodates a large number of meta-frontiers corresponding to a variety of distributions of non-negative random deviations of group frontiers from the meta-frontier, while only two such distributions are (implicitly) allowed by QLP. Third, the CML approach allows for a *simultaneous* estimation of both the local stochastic frontiers and the meta-frontier in the framework of a single data-generating mechanism. The use of copulas, that is, functions representing the extent and nature of the possible stochastic dependence between two and more random variables, allows one to accommodate the possibility of statistical correlation between production shocks associated with the local group frontier and the meta-frontier. It would not be unreasonable to assume, for instance, that countries in a restricted technological environment would have fewer stimuli to produce efficiently and *vice versa*. In order to account for such correlations, we derive a copula-based joint distribution of the two random shocks representing deviations relative to the local group and meta-frontier and find that relaxing the assumption of their stochastic independence results in nontrivial differences in the estimated frontier parameters. Finally, a statistically founded choice can be made between several meta-frontiers when the CML is applied by means of, for example, Akaike information criterion (AIC). This same choice has to be made *ad hoc* (some studies refer to the argument of parsimony) in the QLP framework.

We illustrate our methodology by estimating the meta-frontier parameters in the agricultural sectors of 114 countries for the period of 3 years. We find that the choice of copula and the distributional type of meta-frontier deviation do affect the parameters of the estimated meta-frontiers. Our empirical results also suggest that deviations from the group and meta-frontiers are positively dependent, a conclusion which is impossible to reach within the QLP framework.

This paper is organised as follows. In Section 2, we review the existing approaches to the estimation of meta- and local group frontiers and describe our copula-based approach to the simultaneous estimation of group and meta-frontiers. Section 3 briefly describes the data, and Section 4 presents an empirical illustration of our methodology. Section 5 concludes.

2. Estimation methodology

2.1. Stochastic frontier framework

The original model behind the stochastic frontier analysis (SFA) was first introduced in Aigner *et al*. (1977) and Meeusen and van den Broeck (1977). This model decomposed deviations of the observed output levels from the estimated stochastic production frontier into the part attributable to the uncontrollable factors (e.g. weather), and a non-negative part caused by the producers' inefficient behaviour:

$$\ln(y_i) = (\ln x_i)'\beta + v_i - u_i = (\ln x_i)'\beta + \varepsilon_i, i = 1, 2, \ldots, N \tag{1}$$

where y_i and x_i are the output level and input vectors of firm i, respectively. The symmetric component v_i of the composite error term $\varepsilon_i = v_i - u_i$ is distributed i.i.d. normally with mean zero and variance σ_v^2, while the non-negative term $u_i \geq 0$ that accounts for inefficiency is i.i.d. as well with mean μ and variance σ_u^2. Battese and Coelli (1988) suggest measuring inefficiency levels as a ratio of the observed output level to that on the stochastic production frontier:

$$TE_i = E\left[\frac{\exp(x_i\beta + v_i - u_i)}{\exp(x_i\beta + v_i)}\right] = \exp[-u_i] \in [0, 1] \tag{2}$$

2.2. Meta-frontier Framework

Hayami and Ruttan (1971) appear to be the first introducing the concept of a meta-production function defined as an 'envelope of the neoclassical production frontiers'. Battese and Rao (2002), Battese *et al.* (2004), and O'Donnell *et al.* (2008) develop a meta-frontier model that assumes a specific functional form for the meta-frontier that overarches all group frontiers at the observed combinations of inputs:

$$\begin{cases} \ln(y_i^k) = (\ln x_i^k)'\beta^* \\ (\ln x_i^k)'\beta^* \geq (\ln x_i^k)'\beta^k, k = 1..K, i = 1..N_k \end{cases} \tag{3}$$

where K is the number of such groups, N_k is the number of observations in each group, and the meta-frontier $\left(\ln x_i^k\right)'\beta^*$ overarches the group frontiers $\left(\ln x_i^k\right)'\beta^k$ estimated according to (1) for all observations.

To measure the extent to which the technological environment in a country group is different from the unrestricted one, O'Donnell *et al.* (2008) suggest looking at the meta-technology ratios (MTR) computed as follows:

$$MTR_i^k = \frac{\left(\ln x_i^k\right)'\beta^k}{\left(\ln x_i^k\right)'\beta^*} \tag{4}$$

The level of total technical efficiency $\hat{T}E_i$ of a decision-making unit i with respect to the meta-frontier can be decomposed as follows:

$$\hat{T}E_i^k = TE_i^k \times MTR_i^k \tag{5}$$

2.3. Estimation of Meta-frontier parameters by linear and quadratic programming

Battese *et al.* (2004) offer two algebraic procedures to estimate the parameters in (3):

$$\begin{cases} \underset{\beta^*}{Min} \sum_{k=1}^{k} \sum_{i=1}^{N_k} \left| \left(\ln x_i^k\right)'\beta^* - \left(\ln x_i^k\right)'\hat{\beta}^k \right| \\ \left(\ln x_i^k\right)'\beta^* \geq \left(\ln x_i^k\right)'\hat{\beta}^k \end{cases} \tag{6}$$

and

$$\begin{cases} \underset{\beta^*}{Min} \sum_{k=1}^{K} \sum_{i=1}^{N_k} \left(\left(\ln x_i^k\right)'\beta^* - \left(\ln x_i^k\right)'\hat{\beta}^k \right)^2 \\ (\ln x_i)'\beta^* \geq (\ln x_i)'\hat{\beta}^k \end{cases} \tag{7}$$

where $\hat{\beta}^k$ are the *estimated* parameters of the individual group frontiers. The latter can be estimated by, for example, STATA or FRONTIER 4.1. The solutions to linear and quadratic programming problems (6) and (7), respectively, can be obtained by, for example, SAS or LIMDEP.

The two-step procedures in (6) and (7) are producing estimates of the individual frontier parameters β^k in the first step, employing them to estimate meta-frontier parameters β^* in the second step. However, the argument by Schmidt (1976) implies that this second step is equivalent to finding the parameters of a constrained likelihood function based on the following model:

$$\begin{cases} \left(\ln x_i^k\right)'\hat{\beta}^k = \left(\ln x_i^k\right)'\beta^* - \eta_i^k, i = 1..N_k, k = 1..K \\ \eta_i^k \geq 0 \\ \eta_i^k \sim f_\eta(\cdot) \end{cases} \tag{8}$$

where $f_\eta(\cdot)$ is an exponential or half-normal pdf. In case, $f_\eta(\cdot)$ is exponential, (8) is equivalent to (6), while in the half-normal case, (8) is equivalent to (7).

Indeed, in case η_i^k is distributed exponentially, the log-likelihood function corresponding to (8) is $\ln L = -\left(\sum\limits_{k=1}^{K} N_k\right) \times \ln \sigma_\eta - \frac{1}{\sigma_\eta}\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{N_k}\left(\left(\ln x_i^k\right)'\beta^* - \hat{y}_i^k\right)$, using the parameterisation $f_\eta(\eta) = \frac{1}{\sigma_\eta}e^{-\frac{\eta}{\sigma_\eta}}$ and denoting $\hat{y}_i^k = \left(\ln x_i^k\right)'\hat{\beta}^k$. The first-order conditions imply that the maximum likelihood estimator for the variance of η_i^k is equal to $\hat{\sigma}_\eta = \frac{1}{T}\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{N_k}\left(\left(\ln x_i^k\right)'\beta^* - \hat{y}_i^k\right)$, where $T = \left(\sum\limits_{k=1}^{K} N_k\right)$. The concentrated maximum log-likelihood function is then $\ln L^c = -T - T \times \ln\left(\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{N_k}\left(\left(\ln x_i^k\right)'\beta^* - \hat{y}_i^k\right)\right)$ subject to the constraints $\left(\ln x_i^k\right)'\beta^* - \hat{y}_i^k \geq 0$. Maximising $\ln L^c$ is equivalent to solving problem (6) suggested by Battese *et al*. (2004). The equivalence of problem (7) to the problem of solving (8) under the assumption of half-normal $f_\eta(\cdot)$ is proven analogously and is omitted here for the sake of brevity. It follows that the algebraic procedures (6) and (7) are implicitly based on statistical assumptions on the distribution of η_i^k.

2.4. Simultaneous estimation of individual and group frontiers

As argued above, the second step in Battese *et al*. (2004) is equivalent to estimating the deterministic frontier model in (8) by CML assuming either exponential or half-normal η_i^k. In this study, we expand the set of potential meta-frontiers by allowing for more distributions of η_i^k since the latter does not have to be necessarily exponential or half-normal. We are also offering to estimate both group frontier parameters β^k and the meta-frontier coefficients β^* *simultaneously in one step*, so that the coefficients of both the group frontiers and of the meta-frontier are derived from one single data-generating mechanism:

$$\begin{cases} (\ln x_i^k)'\beta^k = (\ln x_i^k)'\beta^* - \eta_i^k \\ \eta_i^k \geq 0, \eta_i^k \sim f_\eta(\mu_{ik}^\eta, \sigma_k^\eta) \\ \ln(y_i^k) = (\ln x_i^k)'\beta^k + v_i^k - u_i^k, \\ u_i^k \geq 0, u_i^k \sim f_u(\mu_{ik}^u, \sigma_k^u), v_i^k \sim f_v(0, \sigma_k^v) = N(0, \sigma_k^v) \\ i = 1, 2, \ldots, N_k, k = 1..K \end{cases} \tag{9}$$

where $f_u(\mu_{ik}^u, \sigma_k^u)$ is the probability density of a non-negative stochastic component u_i^k, and $f_\eta(\mu_{ik}^\eta, \sigma_k^\eta)$ is the pdf of a non-negative deviation η_i^k of the individual frontiers from the meta-frontier.

The basic idea behind the simultaneous estimation of (9) is that one can associate a three-dimensional random variable (v_i^k, u_i^k, η_i^k) with each observed output level y_i^k. The exogenous shock v_i^k (e.g. bad weather) is outside of control by the producer, the second component $u_i^k \geq 0$ is representing the extent of productive efficiency, and $\eta_i^k \geq 0$ is deviation of the group frontier from the meta-frontier. Assuming we know the joint distribution $g(v_i^k, u_i^k, \eta_i^k)$, both the group frontiers and the meta-frontier parameters in (9) can be estimated simultaneously by maximising the following log-likelihood function subject to non-negativity constraints on η_i^k:

$$\begin{cases} \underset{\beta^k, \beta^*}{Max} \sum_{k=1}^{K} \sum_{i=1}^{N_k} \ln g(v_i^k, u_i^k, \eta_i^k) \\ \eta_i^k \geq 0, i = 1..N_k, k = 1..K \end{cases} \tag{10}$$

While it is standard in the stochastic frontier literature to assume independence between u_i^k and v_i^k (i.e. efforts are independent of the weather conditions), it appears reasonable to assume independence between η_i^k and v_i^k, too. That is, exogenous shocks such as bad weather are unlikely to be correlated with the extent to which a country group's technological environment is restricted. However, restricted economic environment can take its toll on productive efficiency since economic returns to efficient behaviour may be lower in a technologically restricted environment. In this case, we should observe a positive stochastic dependence between η_i^k and u_i^k.

Another advantage of (9) over the two-step procedure is that the simultaneous estimation is more consistent in that it is based on the true values of the group frontier parameters β^k rather than their estimated values $\hat{\beta}^k$. Since the distribution of $\hat{\beta}^k$ is in general unknown, applying the two-step procedure may result in biased estimates of the meta-frontier coefficients.

2.5. Dealing with the Potential Identification Problem

The model in (9) can be further consolidated in the following fashion:

$$\ln\left(y_i^k\right) = \left(\ln x_i^k\right)'\beta^* - \eta_i^k + v_i^k - u_i^k \tag{11}$$

with the potential identification problem regarding η_i^k and u_i^k.

Since we are assuming that v_i^k is independent of both η_i^k and u_i^k, the joint distribution density in (10) can be rendered as $g\left(v_i^k, u_i^k, \eta_i^k\right) = f_v\left(v_i^k\right)h\left(u_i^k, \eta_i^k\right)$ where $h\left(u_i^k, \eta_i^k\right)$ is the joint distribution density of u_i^k and η_i^k. The identification problem will be resolved once $f_v\left(v_i^k\right)h\left(u_i^k, \eta_i^k\right)$ is shown to be a function of the parameters and observed input and output levels in model (9). It follows directly from the first equation in (9) that $\eta_i^k = \left(\ln x_i^k\right)'\beta^* - \left(\ln x_i^k\right)'\beta^k$. The third equation in (9) implies that the symmetric part of the composite error term in the group frontier equation is equal to $v_i^k = \varepsilon_i^k + u_i^k = \ln\left(y_i^k\right) - \left(\ln x_i^k\right)'\beta^k + u_i^k$. As suggested by Greene (2008, p. 177), the mean of the group inefficiency term u_i^k conditional on the $\varepsilon_i^k = v_i^k - u_i^k$ can be estimated as follows:

$$E\left(u_i^k \middle| \varepsilon_i^k\right) = \frac{\int_0^\infty u_i^k f_u\left(u_i^k\right) f_v\left(\varepsilon_i^k + u_i^k\right) du_i^k}{\int_0^\infty f_u\left(u_i^k\right) f_v\left(\varepsilon_i^k + u_i^k\right) du_i^k} \tag{12}$$

Since $\varepsilon_i^k = v_i^k - u_i^k = \ln\left(y_i^k\right) - \left(\ln x_i^k\right)'\beta^k$ is a function of the model's parameters β^k and observations y_i^k, x_i^k the conditional mean $E\left(u_i^k \middle| \varepsilon_i^k\right)$ in (12) is a function of parameters and observations as well. Greene (2008) shows that in case u_i^k is distributed exponentially, its conditional expectation can be rewritten as follows:

$$E\left(u_i^k \middle| \varepsilon_i^k\right) = z_i^k + \sigma_k^v \frac{\varphi\left(z_i^k / \sigma_k^v\right)}{\Phi\left(z_i^k / \sigma_k^v\right)}, z_i^k = \varepsilon_i^k - \frac{\sigma_k^v}{\sigma_k^u} = \ln\left(y_i^k\right) - \left(\ln x_i^k\right)'\beta^k - \frac{\sigma_k^v}{\sigma_k^u} \tag{13}$$

Throughout this paper, we will be assuming without loss of generality the exponential distribution for u_i^k. For more general, one-sided distributions of u_i^k the integrals in (12) have to be computed either numerically or by simulation.

Given the discussion above, we can approximate the joint distribution density in (10) as follows:

$$g\left(v_i^k, u_i^k, \eta_i^k\right) = f_v\left(v_i^k\right)h\left(u_i^k, \eta_i^k\right) \approx f_v\left(\varepsilon_i^k + E\left(u_i^k \middle| \varepsilon_i^k\right)\right)h\left(E\left(u_i^k \middle| \varepsilon_i^k\right), \eta_i^k\right) \tag{14}$$

The 'approximately equal' sign emphasises the fact that, since the exact estimates of u_i^k are unavailable, we approximate for u_i^k using conditional expectations $E\left(u_i^k \middle| \varepsilon_i^k\right)$. The right-hand side of (14) is a function of the meta-frontier model parameters and observed input and output levels alone. Indeed, as demonstrated above, $\varepsilon_i^k = \ln\left(y_i^k\right) - \left(\ln x_i^k\right)'\beta^k$,

$\eta_i^k = \left(\ln x_i^k\right)'\beta^* - \left(\ln x_i^k\right)'\beta^k$ and

$$E\left(u_i^k \middle| \varepsilon_i^k\right) = z_i^k + \sigma_k^v \frac{\varphi\left(z_i^k/\sigma_k^v\right)}{\Phi\left(z_i^k/\sigma_k^v\right)}, z_i^k = \ln\left(y_i^k\right) - \left(\ln x_i^k\right)'\beta^k - \frac{\sigma_k^v}{\sigma_k^u}$$

A Copula Approach to the Derivation of Joint Density of u_i^k and η_i^k

It follows from (14) that the derivation of the joint density of $g\left(v_i^k, u_i^k, \eta_i^k\right)$ boils down to the derivation of the joint density $h\left(u_i^k, \eta_i^k\right)$. To derive the latter, we pursue a copula-based approach. A copula expresses the joint distribution of two or more random variables as a function of the two variables' cumulative density functions called marginal distributions. The fundamental Sklar's (1959) theorem establishes the existence and uniqueness of a copula for any two random variables with continuous marginal distributions. In our case, $H\left(u_i^k, \eta_i^k\right) = C\left(F_u\left(u_i^k\right), F_\eta\left(\eta_i^k\right)\right)$, where $C\left(F_u\left(u_i^k\right), F_\eta\left(\eta_i^k\right)\right)$ is the copula in question, $H\left(u_i^k, \eta_i^k\right)$ is the joint cdf of u_i^k and η_i^k, and $F_u\left(u_i^k\right)$ and $F_\eta\left(\eta_i^k\right)$ are the cdf-s (marginals) of u_i^k and η_i^k, respectively, that may come from different distribution families. It is worthwhile mentioning that, while there is a unique copula for any *particular* joint distribution, the latter does not have to be unique, so that for any two continuous marginal distributions, there may exist a set of copulas linking the two random variables. In terms of the joint *density*, Sklar's theorem implies that $h\left(u_i^k, \eta_i^k\right) = f_u\left(u_i^k\right) f_\eta\left(\eta_i^k\right) c_{12}\left(F_u\left(u_i^k\right),\ F_\eta\left(\eta_i^k\right)\right)$, where $f_u\left(u_i^k\right)$ and $f_\eta\left(\eta_i^k\right)$ are the probability densities of u_i^k and η_i^k, and $c_{12}\left(F_u\left(u_i^k\right), F_\eta\left(\eta_i^k\right)\right)$ is the cross-partial derivative (or copula density) of the copula function $C\left(F_u\left(u_i^k\right), F_\eta\left(\eta_i^k\right)\right)$, see, for example Kumar (2010). Copula functions are often parameterised to reflect the extent to which the two random variables are stochastically dependent.

Hoeffding (1940) and Fréchet (1951) demonstrated that (in the bivariate case) the following must hold:

$$\max[F(x_1) + G(x_2) - 1] \leq C(F(x_1), G(x_2)) \leq \min[F(x_1), G(x_2)] \qquad (15)$$

where x_1 and x_2 are two random variables with marginal distribution cdf-s $F(x_1)$ and $G(x_2)$, respectively. The left- and right-hand side limits of the inequality in (15) are called lower and upper Frechet–Hoeffding bounds, respectively.

Trivedi and Zimmer (2005) in their thorough overview of copulas and related applications stress that the estimation of copulas allows one to verify the property of positive or negative stochastic dependence between two random variables. The two random variables are said to be positively (negatively) dependent in case the copula associated with their joint distribution is achieving the upper (lower) Frechet–Hoeffding bound (Trivedi and Zimmer 2005). In the context of this study, a positive association between u_i^k and η_i^k would mean that larger values of η_i^k are likely to be associated with the larger values of u_i^k as well. Economically speaking, that would mean that in a technologically restricted environment (large η_i^k), producers *within the region* are not likely to be efficient either (larger u_i^k). For instance, in case of Farlie–Gumbel–Morgenstern (FGM) copula given by $C(F(x_1), G(x_2)) = F(x_1)G(x_2)(1 + \rho(1 - F(x_1))(1 - G(x_2)))$, the two variables x_1 and x_2 are positively dependent if the copula parameter ρ is positive, while its negative value implies a negative association. It is easy to see that $\rho = 0$ implies stochastic independence since in this case, the FGM copula collapses to $C(F(x_1), G(x_2)) = F(x_1)G(x_2)$, which corresponds to the independence case. Thus, the copula approach to deriving the joint density function $h\left(u_i^k, \eta_i^k\right)$ in (14) is not only advantageous in the sense that it allows one to account for the possible stochastic dependence between deviations from the regional and meta-frontier, but also because it makes it possible to infer the nature of this stochastic association.

In this study, we make use of the following five copulas whose exact parameterisation can be found, for example, in a comprehensive survey by Trivedi and Zimmer (2005), Balakrishnan and Lai (2009) or Nelsen (2006): the already mentioned FGM copula, Frank, Clayton, Gumbel and Ali-Mikhail-Haq (AMH) copulas. We chose FGM, Frank and AMH copulas since these allow us to model both positive and negative stochastic dependence. Since our empirical results strongly suggest a positive dependence between u_i^k and η_i^k, we add Clayton and Gumbel copulas to our analysis because they are modelling positive dependence.

3. Data

We downloaded our data from the Food and Agriculture Organization's website, faostat.fao.org. The output variable is defined as net production of crops and livestock, measured in thousands of constant international dollars for the period of 2005–2007. The three input factors are labour, land and capital stock. Labour is total economically active population in agriculture, in thousand people. Land includes arable land, land under permanent crops and land under permanent pasture, in thousand hectares. Capital stock is measured in constant prices of 2005 as the market value of physical assets employed in agriculture.

There are 114 countries in our panel, covering the 3 years between 2005 and 2007, resulting in 342 observations. We divide these countries into the following six groups: Advanced Economies, Sub-Saharan African, Middle Eastern and North African, Latin American, South Asian, and East Asian and Pacific countries, the division being based on the World Bank regional classification. The exact country groupings and summary statistics are provided in Tables A1 and Table A2 in the Appendix I, respectively. The use of this data set was inspired by O'Donnell *et al.* (2008).

4. Empirical Application

4.1. Empirical Specifications

We employ the flexible translog functional form for each country group in order to model the local group stochastic production frontiers:

$$\begin{aligned}\ln y_i^k = {} & \beta_0^k + \beta_N^k \ln N_i^k + \beta_L^k \ln L_i^k + \beta_M^k \ln M_i^k + \beta_{NN}^k \left(\ln N_i^k\right)^2 + \beta_{LL}^k \left(\ln L_i^k\right)^2 \\ & + \beta_{MM}^k \left(\ln M_i^k\right)^2 + + \beta_{NL} \ln N_i^k \ln L_i^k + \beta_{NM} \ln N_i^k \ln M_i^k \\ & + \beta_{LM} \ln L_i^k \ln M_i^k + v_i^k - u_i^k \end{aligned} \tag{16}$$

where subscript i refers to the individual countries, while $k = 1.6$ is indexing country groups. The three production factors in (16) are labour N, land L and capital stock M, measured as described in Section 3. The $\beta_{\bullet}^k$ are unknown group frontier parameters to be estimated for each country group.

Random variables v_i^k are symmetric shocks distributed normally with zero mean and variance $\left(\sigma_k^v\right)^2$ and representing the purely random part of each country's output deviation from the deterministic group frontier in a particular year, while random variables u_i^k are representing productive inefficiency. We assume exponentially distributed u_i^k allowing for the group-specific means and variances.

The generalised likelihood test for the null hypothesis that the estimated six group frontiers are statistically indistinguishable from each other confirmed that all group frontiers are, indeed, representing six distinct technological environments. The year dummy was not estimated to be statistically significant in simple production function estimates in either pooled sample or in group regressions, so we assume all production frontiers were stable for the period between 2005 and 2007.

We simultaneously estimate the parameters of group and meta-frontiers by solving the maximisation problem (10), where the joint distribution function $g\left(v_i^k, u_i^k, \eta_i^k\right)$ is derived on the basis of the copula approach discussed in Section 2 according to (14).

Table 1 presents estimates of the meta-frontier parameters computed for five different copulas linking u_i^k and η_i^k mentioned in Section 2. The deviations from meta-frontier η_i^k are assumed to follow an exponential or half-normal distribution to ensure the comparison with the QLP estimation in (6) and (7) corresponding to the assumption of no stochastic dependence between u_i^k and η_i^k. We obtained our estimates by writing a procedure in the R language that employed the 'maxLik' package by Toomet and Henningsen (2012) for CML estimation. Bootstrap errors are reported since, as mentioned in Schmidt (1976), the analytical computation of standard errors is impossible because the range of the dependent variable depends on the estimated parameters of the meta-frontier, that is, $\ln y_i^k \in (-\infty, \left(\ln x_i^k\right)' \beta^*]$. Our bootstrap errors are based on 1000 resamplings with replacement implemented in R.

Table 1 The simultaneous estimates of meta-frontier parameters versus the Linear programming approach

Coefficient	Exponential η_i^k					
	CML estimates, correlated η_i^k and u_i^k					$[Min \sum \lvert\bullet\rvert]$ Linear programming estimates
	FGM	Frank	AMH	Clayton	Gumbel	
C	5.7673 (0.035)	5.7218 (0.023)	5.7578 (0.019)	5.9359 (0.024)	6.1952 (0.019)	5.7460 (0.02)
Ln(N)	4.7902 (0.027)	4.5949 (0.022)	4.6441 (0.020)	4.8738 (0.022)	5.1782 (0.021)	4.6898 (0.019)
Ln(L)	−0.1188 (0.018)	0.0135 (0.015)	−0.0716 (0.014)	−0.7801 (0.055)	0.2444 (0.018)	−0.1523 (0.022)
Ln(M)	0.6688 (0.019)	0.4539 (0.023)	0.5251 (0.014)	0.7724 (0.026)	0.9541 (0.020)	0.5990 (0.015)
Ln^2N	−0.0145 (0.013)	0.0082 (0.011)	0.0006 (0.009)	0.3913 (0.027)	0.5561 (0.091)	−0.0872 (0.014)
Ln^2L	0.1301 (0.009)	0.1184 (0.009)	−0.0014 (0.005)	0.0247 (0.029)	−0.3474 (0.021)	0.0148 (0.005)
Ln^2M	−0.3302 (0.010)	0.1904 (0.007)	0.1305 (0.006)	0.0310 (0.022)	0.4913 (0.036)	0.2092 (0.009)
Ln(N)Ln(L)	0.5772 (0.020)	0.4129 (0.019)	0.2624 (0.011)	0.2907 (0.051)	0.0825 (0.021)	0.1551 (0.013)
Ln(N)Ln(M)	−0.3909 (0.029)	−1.0704 (0.045)	−0.4984 (0.017)	−0.3462 (0.039)	0.0589 (0.019)	0.0432 (0.012)
Ln(L)Ln(M)	0.4122 (0.019)	0.3251 (0.009)	0.1730 (0.004)	0.1780 (0.026)	0.1714 (0.006)	0.3049 (0.009)
$Var(\eta_i^k)$	0.0456	1.2638	0.9034	0.3520	1.9096	0.4417
ρ	0.9305 (0.004)	0.9502 (0.068)	0.8568 (0.023)	12.1784 (0.068)	12.4234 (0.059)	NA

Note: ρ is a copula-specific parameter representing the extent of stochastic dependence between η_i^k and u_i^k. For exact parameterisation and bounds for ρ, see Balakrishnan and Lai (2009) or Nelsen (2006). Bootstrap errors in parentheses.

4.2. Estimation results

As shown by Tables 1 and 2, a meta-frontier that envelops the individual group frontiers most closely in the geometrical sense (i.e. obtained by QLP) is not necessarily the most likely one in the statistical sense. This is an expected result since the QLP methodology ignores stochastic dependence between u_i^k and η_i^k which, however, does not necessarily imply that the QLP estimates of the MTR will be always greater or lower than those obtained under the assumption of stochastic dependence between the two deviation types.

The copula parameter ρ representing the extent of stochastic dependence between u_i^k and η_i^k is estimated to be always statistically significant based on our bootstrap errors. While the FGM, Frank and AMH copulas in principle allow for the negative dependence, the estimated values of ρ are all suggesting strong positive dependence between the one-sided deviations from group and meta-frontiers. This finding implies that in a restricted technological environment (large η_i^k), the productive efficiency of its individual members is likely to be lower (the u_i^k are also large). While intuitively appealing, this inference is hardly possible to make if one assumes stochastic independence between u_i^k and η_i^k. In contrast, the copula approach to simultaneously estimating the meta- and group frontiers makes it possible to infer both the direction and the magnitude of the possible statistical association.

Table 3 below further illustrates how allowing for the stochastic dependence between u_i^k and η_i^k may alter the estimates of MTR. The group of Sub-Saharan African countries is estimated to be falling 70 per cent short of the global meta-frontier assuming stochastically independent deviations from group and meta-frontiers, and exponential η_i^k. However, assuming the two types of deviation are linked by the AMH copula decreases the meta-technology ratio by twenty percentage points. The individual countries are also ranked differently depending on the assumptions on the stochastic independence of u_i^k and η_i^k and the distribution of η_i^k. Thus, the MTR of South Africa is estimated to be 82.9 per cent assuming independent meta- and group frontier deviations and half-normal η_i^k, while it is only 10.5 per cent for the half-normal η_i^k and the FGM copula.

While allowing for stochastic dependence between u_i^k and η_i^k sometimes increases the MTR (e.g. the group of advanced countries for exponential η_i^k and Gumbel copula), the MTR-s estimated under the assumption of stochastic independence are on average higher than their copula-based counterparts by three percentage points.

In Table 4, we present the expanded set of our estimates based on the additional two-one-sided distributions of η_i^k, namely Rayleigh and Weibull. In no one case, do we find evidence of a negative dependence between u_i^k and η_i^k, reinforcing our inference about the less efficient behaviour in the more restricted technological environments. Low bootstrap standard errors on the copula parameter ρ in each estimation case suggest a strong extent of stochastic dependence between deviations from group and meta-frontiers.

Table 2 The Simultaneous Estimates of Meta-Frontier Parameters versus the Quadratic Programming Approach

Coefficient	Half-normal η_i^k					
	CML estimates, correlated η_i^k and u_i^k					$\left[Min \sum (\bullet)^2\right]$ Quadratic programming estimates
	FGM	Frank	AMH	Clayton	Gumbel	
C	5.8511 (0.019)	5.8191 (0.020)	5.8615 (0.020)	5.9053 (0.031)	5.9814 (0.027)	5.7459 (0.019)
Ln(N)	4.8249 (0.016)	4.7958 (0.016)	4.9420 (0.013)	4.8567 (0.029)	4.9202 (0.031)	4.6898 (0.018)
Ln(L)	−0.0420 (0.017)	−0.0630 (0.017)	−0.4078 (0.027)	−0.7028 (0.045)	0.1542 (0.028)	−0.1523 (0.018)
Ln(M)	0.6207 (0.013)	0.5984 (0.013)	0.7641 (0.013)	0.7621 (0.028)	0.7807 (0.024)	0.5990 (0.016)
Ln^2N	−0.1208 (0.015)	−0.1238 (0.015)	0.2246 (0.010)	0.3854 (0.023)	0.1340 (0.026)	−0.0872 (0.016)
Ln^2L	0.0931 (0.008)	0.1118 (0.008)	−0.0228 (0.004)	0.0381 (0.005)	−0.0866 (0.004)	0.0148 (0.005)
Ln^2M	0.0613 (0.006)	0.0683 (0.006)	−0.0128 (0.005)	0.0145 (0.004)	0.0295 (0.003)	0.2092 (0.010)
Ln(N)Ln(L)	0.1023 (0.008)	0.1002 (0.008)	0.1856 (0.010)	0.2606 (0.028)	0.1929 (0.028)	0.1551 (0.013)
Ln(N)Ln(M)	−0.0505 (0.009)	−0.0498 (0.009)	0.0175 (0.011)	−0.3458 (0.018)	0.0208 (0.02)	0.0432 (0.012)
Ln(L)Ln(M)	−0.0251 (0.007)	−0.0341 (0.007)	0.1635 (0.005)	0.1747 (0.006)	0.2014 (0.005)	0.3049 (0.014)
$Var(\eta_i^k)$	0.2119	0.1913	0.3020	0.2075	0.1506	0.665
ρ	0.7338 (0.015)	0.8342 (0.015)	0.7735 (0.026)	12.1447 (0.033)	12.0932 (0.027)	NA

Note: ρ is a copula-specific parameter representing the extent of stochastic dependence between η_i^k and u_i^k. For exact parameterisation and bounds for ρ, see Balakrishnan and Lai (2009) or Nelsen (2006).Bootstrap errors in parentheses.

Table 3 The QLP and CML Estimates of Meta-Technology Ratios, (a) Exponential η_i^k, (b) Half-Normal η_i^k

	Copula type					
	FGM	Frank	AMH	Clayton	Gumbel	Linear programming
(a)						
Country group						
Advanced	0.905 (0.06)	0.897 (0.044)	0.938 (0.04)	0.921 (0.04)	0.961 (0.02)	0.950 (0.03)
East Asia and Pacific	0.707 (0.21)	0.437 (0.29)	0.626 (0.18)	0.933 (0.04)	0.548 (0.14)	0.854 (0.09)
North Africa and Middle East	0.427 (0.39)	0.582 (0.21)	0.730 (0.15)	0.547 (0.36)	0.809 (0.07)	0.636 (0.21)
Sub-Saharan Africa	0.877 (0.06)	0.806 (0.11)	0.500 (0.26)	0.848 (0.15)	0.910 (0.08)	0.698 (0.38)
Latin America	0.597 (0.18)	0.618 (0.17)	0.530 (0.20)	0.569 (0.19)	0.768 (0.09)	0.697 (0.13)
South-East Asia	0.889 (0.06)	0.892 (0.07)	0.900 (0.06)	0.979 (0.02)	0.983 (0.01)	0.925 (0.05)
Countries						
U.S.	0.848 (0.0005)	0.837 (0.0001)	0.899 (0.0003)	0.877 (0.0004)	0.939 (0.0001)	0.905 (0.000)
Japan	0.929 (0.002)	0.997 (0.003)	0.951 (0.0003)	0.900 (0.001)	0.923 (0.0002)	0.934 (0.0003)
China	0.999 (0.0007)	0.998 (0.002)	0.998 (0.002)	0.865 (0.0002)	0.618 (0.001)	0.762 (0.0008)
Indonesia	0.875 (0.002)	0.654 (0.0009)	0.780 (0.001)	0.896 (0.0001)	0.550 (0.001)	0.800 (0.0002)
Egypt	0.171 (0.0008)	0.272 (0.0005)	0.441 (0.001)	0.791 (0.007)	0.844 (0.001)	0.988 (0.011)
Iran	0.024 (0.0001)	0.998 (0.002)	0.989 (0.001)	0.906 (0.0003)	0.802 (0.0003)	0.760 (0.001)
Nigeria	0.920 (0.002)	0.664 (0.003)	0.266 (0.002)	0.977 (0.001)	0.940 (0.0002)	0.999 (0.001)
South Africa	0.898 (0.001)	0.888 (0.002)	0.451 (0.002)	0.880 (0.0002)	0.955 (0.0003)	0.820 (0.001)
Brazil	0.423 (0.000)	0.537 (0.002)	0.304 (0.001)	0.354 (0.001)	0.691 (0.0007)	0.574 (0.0006)
Argentina	0.430 (0.0004)	0.717 (0.002)	0.429 (0.001)	0.413 (0.0001)	0.642 (0.001)	0.600 (0.0003)
India	0.899 (0.004)	0.912 (0.003)	0.920 (0.002)	0.965 (0.001)	0.985 (0.0004)	0.894 (0.0009)
Bangladesh	0.920 (0.0005)	0.897 (0.002)	0.889 (0.002)	0.999 (0.001)	0.982 (0.0001)	0.941 (0.002)

Table 3 (*Continued*)

	Copula Type					
	FGM	Frank	AMH	Clayton	Gumbel	Quadratic programming
(b)						
Country group						
Advanced	0.937 (0.04)	0.928 (0.04)	0.941 (0.03)	0.921 (0.04)	0.944 (0.03)	0.949 (0.03)
East Asia and Pacific	0.896 (0.07)	0.508 (0.18)	0.684 (0.10)	0.933 (0.035)	0.850 (0.05)	0.857 (0.08)
North Africa and Middle East	0.828 (0.13)	0.865 (0.09)	0.735 (0.21)	0.547 (0.36)	0.886 (0.08)	0.624 (0.24)
Sub-Saharan Africa	0.324 (0.30)	0.740 (0.29)	0.780 (0.11)	0.848 (0.15)	0.628 (0.21)	0.695 (0.39)
Latin America	0.394 (0.25)	0.528 (0.24)	0.636 (0.12)	0.569 (0.189)	0.801 (0.10)	0.694 (0.13)
South-East Asia	0.976 (0.02)	0.929 (0.04)	0.983 (0.014)	0.979 (0.016)	0.929 (0.05)	0.931 (0.04)
Countries						
U.S.	0.879 (0.0001)	0.879 (0.0004)	0.915 (0.003)	0.877 (0.0004)	0.916 (0.0003)	0.903 (0.0001)
Japan	0.941 (0.001)	0.878 (0.001)	0.919 (0.0004)	0.900 (0.0008)	0.917 (0.0003)	0.932 (0.0003)
China	0.747 (0.0001)	0.315 (0.001)	0.673 (0.0006)	0.865 (0.0002)	0.860 (0.001)	0.757 (0.0007)
Indonesia	0.836 (0.0005)	0.425 (0.001)	0.661 (0.0008)	0.896 (0.0001)	0.837 (0.001)	0.801 (0.0001)
Egypt	0.707 (0.001)	0.767 (0.004)	0.830 (0.003)	0.791 (0.007)	0.771 (0.002)	0.989 (0.01)
Iran	0.999 (0.001)	0.777 (0.001)	0.956 (0.0005)	0.906 (0.0003)	0.803 (0.001)	0.787 (0.001)
Nigeria	0.168 (0.0005)	0.999 (0.001)	0.706 (0.0005)	0.977 (0.0007)	0.502 (0.0001)	0.999 (0.001)
South Africa	0.105 (0.0002)	0.822 (0.001)	0.710 (0.001)	0.880 (0.0002)	0.453 (0.0002)	0.829 (0.001)
Brazil	0.062 (0.001)	0.256 (0.001)	0.531 (0.0001)	0.354 (0.0006)	0.721 (0.0005)	0.573 (0.0001)
Argentina	0.145 (0.001)	0.185 (0.002)	0.548 (0.0001)	0.413 (0.0001)	0.683 (0.0007)	0.603 (0.0002)
India	0.949 (0.001)	0.912 (0.0004)	0.965 (0.001)	0.965 (0.001)	0.883 (0.0001)	0.905 (0.0006)
Bangladesh	0.991 (0.001)	0.947 (0.002)	0.990 (0.0002)	0.999 (0.0009)	0.942 (0.0007)	0.947 (0.002)

Note: Standard deviations in parentheses. Linear programming estimates are identical with the ones obtained under the assumption of η_i^k distributed exponentially and independent η_i^k and u_i^k.
Quadratic programming estimates are identical with the ones obtained under the assumption of η_i^k distributed half-normally and independent η_i^k and u_i^k.

Table 4 The CML Estimates of Meta-Frontier Parameters and Meta-Technology Ratios for Various Copulas and η_i^k Distributed According to Weibull or Rayleigh Densities

Distribution of η_i^k	Weibull η_i^k					Rayleigh η_i^k				
Copula	FGM	Frank	AMH	Clay-ton	Gum-bel	FGM	Frank	AMH	Clay-ton	Gum-bel
Meta-frontier parameters										
C	6.203 (0.019)	6.210 (0.023)	5.927 (0.027)	6.145 (0.028)	6.098 (0.001)	5.775 (0.022)	5.720 (0.020)	5.758 (0.019)	6.064 (0.030)	6.195 (0.020)
Ln(N)	5.122 (0.021)	5.127 (0.18)	4.878 (0.017)	5.057 (0.029)	5.029 (0.000)	4.796 (0.014)	4.592 (0.021)	4.644 (0.020)	5.014 (0.011)	5.178 (0.010)
Ln(L)	0.165 (0.015)	0.172 (0.011)	−0.012 (0.019)	0.144 (0.026)	0.111 (0.000)	−0.113 (0.013)	0.019 (0.013)	−0.072 (0.014)	0.124 (0.012)	0.244 (0.009)
Ln(M)	0.788 (0.010)	0.784 (0.009)	0.692 (0.011)	0.736 (0.026)	0.728 (0.003)	0.670 (0.011)	0.450 (0.016)	0.525 (0.014)	0.814 (0.014)	0.954 (0.009)
Ln^{2}N	0.137 (0.008)	0.129 (0.004)	0.055 (0.006)	−0.004 (0.023)	0.118 (0.001)	−0.021 (0.008)	0.006 (0.009)	0.001 (0.009)	0.074 (0.015)	0.556 (0.072)
Ln^{2}L	−0.164 (0.007)	−0.194 (0.010)	−0.058 (0.010)	0.072 (0.006)	−0.261 (0.006)	0.137 (0.009)	0.124 (0.007)	−0.001 (0.005)	−0.066 (0.015)	−0.347 (0.010)
Ln^{2}M	0.010 (0.003)	−0.021 (0.005)	0.075 (0.007)	−0.005 (0.010)	0.062 (0.002)	−0.369 (0.014)	0.191 (0.007)	0.130 (0.006)	0.007 (0.028)	0.491 (0.030)
Ln(N)Ln(L)	0.037 (0.008)	0.030 (0.006)	0.038 (0.006)	0.007 (0.020)	0.080 (0.001)	0.574 (0.020)	0.417 (0.015)	0.262 (0.011)	0.044 (0.083)	0.082 (0.021)
Ln(N)Ln(M)	−0.027 (0.009)	−0.031 (0.006)	0.275 (0.013)	−0.087 (0.019)	−0.100 (0.000)	−0.385 (0.014)	−1.083 (0.033)	−0.498 (0.017)	−0.086 (0.021)	0.059 (0.019)
Ln(L)Ln(M)	0.238 (0.005)	0.220 (0.004)	0.185 (0.007)	0.278 (0.010)	0.251 (0.004)	0.420 (0.010)	0.326 (0.007)	0.173 (0.004)	0.124 (0.039)	0.171 (0.004)
σ	1.231 (0.014)	1.211 (0.025)	0.775 (0.028)	1.004 (0.037)	1.017 (0.000)	0.208 (0.029)	0.969 (0.037)	0.951 (0.037)	1.014 (0.023)	1.176 (0.067)
λ	1.204 (0.026)	0.795 (0.018)	0.839 (0.023)	0.991 (0.056)	0.991 (0.001)	NA	NA	NA	NA	NA
Copula parameter ρ	0.692 (0.014)	0.685 (0.008)	0.742 (0.008)	12.223 (0.035)	12.223 (0.002)	0.930 (0.004)	0.922 (0.023)	0.857 (0.023)	12.284 (0.029)	12.423 (0.039)
Meta-Technology Ratios										
Advanced	0.936 (0.03)	0.935 (0.03)	0.944 (0.03)	0.944 (0.03)	0.942 (0.03)	0.871 (0.08)	0.891 (0.04)	0.938 (0.04)	0.921 (0.04)	0.961 (0.02)
East Asia and Pacific	0.881 (0.08)	0.757 (0.08)	0.864 (0.03)	0.864 (0.05)	0.884 (0.07)	0.677 (0.22)	0.432 (0.30)	0.626 (0.18)	0.933 (0.04)	0.548 (0.14)
Northern Africa	0.900 (0.07)	0.912 (0.07)	0.697 (0.13)	0.697 (0.28)	0.784 (0.16)	0.474 (0.30)	0.576 (0.22)	0.730 (0.15)	0.547 (0.36)	0.809 (0.07)
Sub-Saharan Africa	0.803 (0.21)	0.806 (0.21)	0.659 (0.08)	0.659 (0.10)	0.524 (0.25)	0.848 (0.08)	0.818 (0.11)	0.500 (0.26)	0.848 (0.15)	0.910 (0.08)
Latin	0.638 (0.16)	0.637 (0.16)	0.783 (0.19)	0.783 (0.08)	0.731 (0.09)	0.585 (0.19)	0.619 (0.17)	0.530 (0.20)	0.569 (0.19)	0.768 (0.09)
South-Eastern Asia	0.852 (0.08)	0.850 (0.08)	0.951 (0.05)	0.951 (0.03)	0.996 (0.002)	0.883 (0.06)	0.895 (0.07)	0.900 (0.06)	0.979 (0.02)	0.983 (0.01)

Note: σ is a variance-related parameter in case of half-normal, exponential and Rayleigh distributions. In the case of the Weibull distribution, σ is the scale parameter, while λ is the shape parameter. For exact parameterisation, see Balakrishnan and Lai (2009). Copula parameter ρ represents the extent to which η_i^k and u_i^k are stochastically dependent. Bootstrap standard errors are in parentheses for the meta-frontier parameters. Standard deviations are in parentheses for the MTR-s.

O'Donnell *et al*. (2008) appear to be the study most related to ours in that it also applies the meta-frontier approach to the world agricultural data coming from the same FAO database, albeit using different country groupings and covering a different time period. Similarly to the authors who find the average MTR for all countries to be equal to 72.7 per cent, our estimate of the average MTR across all countries and copula specifications is 76.46 per cent. We also estimate China's MTR to be above 99 per cent in three cases (FGM, Frank and AMH copulas for exponential η_i^k), but in all other cases, the Chinese MTR is estimated to be lower. There are similarities with other studies of meta-frontiers as well. For instance, similarly to Battese *et al*. (2004) who estimate a meta-frontier for the Indonesian regions for the garment industry, we find the estimated MTR-s vary considerably across regions within the same specification of η_i^k, as well as across different specifications. In addition, we also find variation in our estimates depending on the type of copula.

We believe an important insight obtained in this paper that is impossible to gain if the issue of stochastic correlation between u_i^k and η_i^k is ignored is the strong evidence of positive dependence between deviations from meta- and group frontiers. This positive dependence implies that a higher-quality, less restrictive technological environment in the region is positively associated with the more efficient production practices within the region. However, this finding immediately poses the question of the direction of causality: is it that a high-quality technological environment that creates incentives for the individual decision makers to be more efficient, or is it that the individual efforts to produce more efficiently that eventually weaken the restrictions on the technological environment as a whole? While it is impossible to answer this question within the scope of this study, we believe it is an important issue to analyse in the future.

In general, both the magnitudes and the ranking of meta-frontier parameters and the MTR differ depending on the assumptions regarding the distribution of η_i^k and the copula density $c_{12}\left(F_u\left(u_i^k\right), F_\eta\left(\eta_i^k\right)\right)$, which raises the problem of choosing the 'best' meta-frontier. We use Akaike (1973) information criterion (AIC) to choose among the variety of meta-frontiers. The AIC criterion statistic is computed as $-2 \ln L + 2k_f$, where k_f is the number of degrees of freedom equal to the number of the estimated parameters. The criterion 'favors' the model with the lowest value of the AIC statistic.

According to Table 5, the meta-frontier model based on the Gumbel copula assuming a Weibull distribution for η_i^k vastly outperforms all the other alternatives. Relative likelihood values reported in parentheses are estimated according to Burnham and Anderson (2002) as $\exp\left\{\left(AIC_{\min} - AIC_j\right)/2\right\}$ where $\mathrm{AIC}_{\min}$ and AIC_j are the AIC values of the model with the minimum AIC and the AIC for model j, respectively. Conceptually, relative likelihood can be thought of as the probability of model j being as good in the statistical sense as the one corresponding to the minimum value of the AIC. We

Table 5 Akaike Criterion Statistics for Alternative Meta-Frontier Specifications

Distribution of η_i^k	Copulas					No correlation
	FGM	Frank	AMH	Clayton	Gumbel	
Half-normal	−215.16	−255.39	−440.41	−476.70	−477.96	−363.91
	(0.00%)	(0.00%)	(0.00%)	(0.14%)	(0.25%)	(0.00%)
Exponential	−299.64	−283.89	−381.82	−484.80	−220.27	−383.53
	(0.00%)	(0.00%)	(0.00%)	(7.79%)	(0.00%)	(0.00%)
Weibull	−424.85	−427.95	−243.56	−471.53	−489.91	−364.04
	(0.00%)	(0.00%)	(0.00%)	(0.01%)	(100.00%)	(0.00%)
Rayleigh	−307.00	−283.89	−288.82	−475.80	−221.82	−426.45
	(0.00%)	(0.00%)	(0.00%)	(0.09%)	(0.00%)	(0.00%)

Note: Relative likelihood values in parentheses.

interpret the estimates in Table 5 to strongly suggest that in our context, it is best to model meta-frontiers assuming a Weibull distribution for η_i^k and the Gumbel copula linking u_i^k and η_i^k.

5. Conclusions

In this study, we have suggested a copula-based approach to the simultaneous estimation of meta- and group frontiers, which is different from the two-step procedure by Battese *et al.* (2004). Our approach is based upon a single data-generating mechanism for the parameters of both meta- and group frontiers, which in our view makes it preferable to the two-stage approach. In addition, we expand the set of possible meta-frontiers by allowing for different one-sided distributional assumptions on the deviations from meta-frontiers and for the different nature of stochastic dependence between η_i^k and group frontier deviations u_i^k captured by copulas.

We believe it is important to account for the stochastic dependence between u_i^k and η_i^k for several reasons. First, we find consistent evidence in favour of stochastic dependence between deviations from group and meta-frontiers u_i^k and η_i^k. Second, the estimates of both the meta-frontier parameters and those of the MTR do differ depending on whether we assume u_i^k and η_i^k to be stochastically dependent both in magnitude and in terms of rankings. Third, ignoring stochastic dependence between the two types of deviation leads to overlooking an important characteristic of the economic environment related to the interaction between the extent to which a technological environment is restricted in a group of countries and the efficient behaviour within the country group. Applying our estimation methodology to the data on world agricultural production, we find that u_i^k and η_i^k are strongly and positively dependent, implying that productive efficiency levels of individual countries will likely be lower in the more restricted technological environments. This inference is hardly possible to make in case one ignored the possibility of stochastic dependence between u_i^k and η_i^k.

Finally, our approach allows for an educated choice between the multitude of meta-frontiers based on the values of the likelihood functions corresponding to different meta-frontier model specifications. We find that the meta-frontier model based on η_i^k following Weibull distribution and the Gumbel copula linking u_i^k and η_i^k is best describing our data set according to the AIC.

References

Aigner, D., Lovell, C.A.K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* 6, 21–37.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, in Petrov, B.N. and Csaki, F. (eds), *2nd International Symposium on Information Theory*. Akadémiai Kiadó, Budapest, pp. 267–281.

Balakrishnan, N. and Lai, C.H.D. (2009). *Continuous Bivariate Distributions*. Springer, New York, pp. 33–60.

Battese, G.E. and Coelli, T. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* 38(3), 387–399.

Battese, G.E. and Rao, D.S.P. (2002). Technology gap, efficiency, and a stochastic metafrontier function, *International Journal of Business and Economics* 1(2), 1–7.

Battese, G.E., Rao, D.S.P. and O'Donnell, C.J. (2004). A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies, *Journal of Productivity Analysis* 21, 91–103.

Burnham, K.P. and Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer, New York.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données, *Annales de l'Université de Lyon Section A*14, 53–77.

Greene, W.H. (2008). The econometric approach to efficiency analysis, in Fried, H.O., Knox, K.L. and Schmidt, S.S. (eds), *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, New York, pp. 92–250.

Hayami, Y. and Ruttan, V.W. (1971). *Agricultural Development: An International Perspective*. Johns Hopkins University Press, Baltimore.

Hoeffding, W. (1940). Masstabinvariante Korrelationstheorie, *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universitat Berlin* 5, 179–233.

Kumar, P. (2010). Probability distributions and estimation of Ali-Mikhail-Haq Copula, *Applied Mathematical Sciences* 4(14), 657–666.

Meeusen, W. and van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review* 18, 435–444.

Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd edn. Springer, New York.

O'Donnell, C.J., Rao, D.S.P. and Battese, G.E. (2008). Metafrontier frameworks for the study of firm-level efficiencies and technology ratios, *Empirical Economics* 34, 231–255.

Schmidt, P. (1976). On the statistical estimation of the frontier parametric production functions, *Review of Economics and Statistics* 58, 238–239.

Sklar, A. (1959). Fonctions de Répartition à *n* Dimensions et Leurs Marges, *Publications de l'Institut de Statistique de l'Université de Paris* 8, 229–231.

Toomet, O. and Henningsen, A. (2012). Package maxLik. Available from URL .http://www.maxLik.org, [accessed 10 April 2012].

Trivedi, P.K. and Zimmer, D.M. (2005). Copula modeling: an introduction for practitioners, *Foundations and Trends in Econometrics* 1(1), 1–111.

Appendix: Country Listings and Summary Statistics

Table A1 Groups of Countries for the Empirical Analysis

Advanced Economies	South Asia	Sub-Saharan Africa	Middle East and North Africa	East Asia and Pacific	Latin America
Australia	Afghanistan	Angola	Algeria	Cambodia	Antigua and Barbuda
Austria	Bangladesh	Botswana	Djibouti	China	Argentina
Belgium	Bhutan	Burkina Faso	Egypt	Indonesia	Bahamas
Canada	India	Burundi	Iran	Laos	Barbados
Denmark	Nepal	Cameroon	Iraq	Malaysia	Belize
Finland	Pakistan	Chad	Israel	Mongolia	Bolivia
France	Sri Lanka	Cote d'Ivoire	Jordan	Myanmar	Brazil
Germany		Ghana	Lebanon	Papua New Guinea	British Virgin Islands
Greece		Guinea	Libya	Philippines	Chile
Ireland		Guinea -Bissau	Morocco	Thailand	Colombia
Italy		Kenya	Syria	East Timor	Costa Rica
Japan		Lesotho	Tunisia	Viet Nam	Cuba
Mexico		Madagascar	Yemen		Dominica
Netherlands		Malawi			Dominican Republic
New Zealand		Mali			Ecuador
Norway		Mauritania			El Salvador
Portugal		Mozambique			Grenada
Republic of Korea		Niger			Guadeloupe
Spain		Nigeria			Guatemala
Sweden		Rwanda			Guyana, Haiti
Switzerland		Senegal			Honduras, Jamaica
Turkey		Sierra Leone			Nicaragua, Panama
United Kingdom		Somalia			Paraguay
United States		South Africa			Peru
		Sudan			Puerto Rico
		Uganda			Trinidad and Tobago
		Zimbabwe			Uruguay, Venezuela

Table A2 Summary Statistics

	Advanced economies		South Asia		Sub-Saharan Africa	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Output (mn I$)	24,797	43,866	36,695	67,419	3800	6586
Labour (thousand)	1277	2266	47,631	88,719	4122	3257
Land (mn Ha)	50.49	116	37.34	61	25.72	25
Capital (mn I$)	77,959	120,468	81,893	118,229	10,258	12,571
	Middle East and North Africa		East Asia and Pacific		Latin America	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Output (mn I$)	6078	7705	52,125	127,373	7140	21,493
Labour (thousand)	1881	2260	53,713	138,759	1075	2219
Land (mn Ha)	15.27	15.51	62.89	144	19.50	52.20
Capital (mn I$)	19,031	21,551	68,423	145,617	19,002	40,947

Note: I$ is international dollars, base 2005; we are employing 342 observations in all.