Routledge
Taylor & Francis Group

Agrekon

Agricultural Economics Research, Policy and Practice in Southern Africa

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ragr20

Technological differences in South African sheep production: a stochastic meta-frontier analysis

Nicolette Matthews, Beatrice Conradie & Jenifer Piesse

To cite this article: Nicolette Matthews, Beatrice Conradie & Jenifer Piesse (2023) Technological differences in South African sheep production: a stochastic meta-frontier analysis, Agrekon, 62:1, 19-30, DOI: 10.1080/03031853.2022.2149577

To link to this article: https://doi.org/10.1080/03031853.2022.2149577

Published online: 27 Jan 2023.

AGREKON
2023, VOL. 62, NO. 1, 19–30
https://doi.org/10.1080/03031853.2022.2149577

Technological differences in South African sheep production: a stochastic meta-frontier analysis

Nicolette Matthews [a], Beatrice Conradie [b] and Jenifer Piesse [c]

[a]Agricultural Economics, University of the Free State, Bloemfontein, South Africa; [b]School of Economics, University of Cape Town, Cape Town, South Africa; [c]Agricultural Economics, Stellenbosch University, Stellenbosch, South Africa

ABSTRACT

This study compared four South African sheep producing districts relative to each other and a common metafrontier to analyse within and between group efficiency and explored what could be learnt from this technique compared to simple frontiers. A sample was compiled from sources that were previously successfully used in local benchmarking exercises, and despite very modest sample sizes at the group level and minimal information on how groups differ, the group models performed adequately while the meta-model performed very well. The results revealed that while within group performances were comparable across districts, there were huge differences in between group performance. These differences are partly attributable to natural resource endowments, but institutional arrangements also contribute significantly to local success. This suggests that to achieve rural regeneration public–private partnerships are necessary to address this issue. State support is insufficient and producer organisations have a major role in promoting institutional innovation.

ARTICLE HISTORY
Received 31 October 2022
Accepted 3 November 2022

KEYWORDS
Stochastic metafrontier analysis; technological gaps; convergence; innovation; small stock

1. Introduction

There are many examples in the literature where productivity efficiency is measured across groups. These can include countries, provinces or states as well as enterprises that have organisational differences, such as private versus state owned or cooperatives. But frequently these give spurious results as the groups are simply too disparate to be estimated using a single frontier. Equally, estimating separate frontiers and comparing the derived efficiencies is also unjustified as the groups may not have access to the same technology.

To overcome these difficulties Hayami and Ruttan (1971) proposed the concept of the meta-production function. This assumes that in theory, all producers have access to the same set of technologies but will adopt those that are available to them given their unique circumstances. There may be differences in input prices, environment, climate, natural resource endowments as well as financial and regulatory constraints. The presence of one or more of these constraints may limit access to the best practice technology and results in a gap between potential and actual production. Therefore, the difference between the metafrontier and the group-specific frontiers must be measured and this is known as the technology gap ratio (TGR). See Kawagoe, Hayami, and Ruttan (1985) for an early empirical estimation.

More recently, models to conduct metafrontier analysis (MFA) have been proposed by Battese, Rao, and O'Donnell (2004) and an application to regional performance in agriculture is by O'Donnell, Rao, and Battese (2008). Metafrontiers can be derived semi-parametrically or fully parametrically.

CONTACT Beatrice Conradie beatrice.conradie@uct.ac.za

Battese, Rao, and O'Donnell (2004) and O'Donnell, Rao, and Battese (2008) used semi-parametric MFA models, which estimate group stochastic frontier models introduced by Battese and Coelli's (1995) in the first stage and then apply mathematical programming to derive a closely fitting metafrontier that minimises deviations from the group frontiers. In the local literature variants of this technique have been used to compare smallholder cattle production across districts in Botswana (Temoso et al., 2015), the gap between smallholder and commercial tomato production in Mpumalanga, South Africa (Gwebu and Matthews 2018) and smallholder sheep farming in adjacent districts in the Free State, South Africa (Nyam, Matthews, and Bahta 2020).

The disadvantage of a semi-parametric analysis is that the metafrontier does not result in parameters with known statistical properties from which inferences can be made. Therefore, to avoid this problem Huang, Huang, and Liu (2014) derived a measure of the technology gap ratio (TGR) that can be obtained directly from estimating a solely parametric metafrontier. These two approaches define TGRs differently, which can cause confusion. See Section 2 below and Huang, Huang, and Liu (2014) for the derivation of the stochastic TGR measure. Huang, Huang, and Liu (2014) repeated O'Donnell's study to demonstrate the effects of using a parametric second stage, and as would be expected, found that the semi-parametric approach leads to smaller gaps than the fully parametric method. This is because all deviations are treated as part of the TGR in the non-parametric approach, whilst the parametric approach allows for a stochastic measurement error.

Only one example of the Huang, Huang, and Liu (2014) approach is found in the local literature. Ng'ombe (2017) estimated a stochastic metafrontier model of technological diversity in smallholder maize farming in Zambia. The functional form of the frontier was specified as a four-input translog function which is a flexible function form that imposes few restrictions itself and, in this case, included ten inefficiency effects. In the MFA function groups were formed by province. Lusaka Province had the smallest number of observations ($n = 312$) and these degrees of freedom produced a set of group frontiers with a good fit and an even better metafrontier. The inefficiency model in the metafrontier included just three dummy variables capturing regional differences in expected rainfall. The results found significant technological diversity in maize production in Zambia, but the efficiency ordering did not simply follow rainfall patterns as anticipated.

This study uses a stochastic metafrontier model with four groups of sheep farmers (three in the Karoo and one in the Overberg in the Western Cape). This formulation has far fewer observations and potential explanations of farm-level inefficiencies within groups than Ng'ombe (2017) had but the aim is to identify something potentially important for parametric productivity analysis; to find potentially elusive technological heterogeneity in a single industry in one region of one country. The sector under consideration is sheep production and the scope the of study is limited to the Northern and Western Cape Provinces of South Africa. If pooling is rejected in this example, it is likely to be rejected for all composite samples, and as a result metafrontier analysis will become the norm in benchmarking studies of agriculture, which is not currently the case.

2. Data and methods

2.1 Deriving SMF estimators and technology gap measures

Step 1 of the stochastic metafrontier (SMF) analysis begins with estimating group frontiers specified in the usual manner for every production region, j, for every farm i for each period t as follows:

$$Y_{ijt} = f_t^j(X_{ijt})e^{v_{ijt}-u_{ijt}}, \qquad j = 1, 2, \ldots, J; \; i = 1, 2, \ldots, I; \; t = 1, 2, \ldots, T \tag{1}$$

where: Y_{ijt} is the output of farm i in group j at time t X_{ijt} are input vectors related to farm i in group j at time t $f_t^j(\cdot)$ is the functional form of the production frontier to be estimated for the jth group v_{ijt} is statistical noise u_{ijt} is technical inefficiency

The v_{ijt} are assumed to be independently and identically distributed as $N(0, \sigma_v^{j2})$ and these errors are independent of u_{ijt}. The inefficiency term, u_{ijt}, follows a truncated-normal distribution,

$N(\mu^j(Z_{ijt}), \sigma_u^{j2}(Z_{ijt}))$, where Z_{ijt} are exogenous variables that explain the level of inefficiency observed, and $\mu^j(Z_{ijt})$ is the mode of the inefficiency term.

A set of group technical efficiency estimates, TE^j_{it}, is obtained from each group's production frontier:

$$TE^j_{it} = \frac{Y_{ijt}}{f^j_t(X_{ijt})e^{v^j_{it}}} = e^{u_{ijt}} \tag{2}$$

In stage 2 a metafrontier is estimated by regressing fitted values of the group frontiers, $\hat{f}^j_t(X_{ijt})$, on the original inputs according to regression 3:

$$\hat{f}^j_t(X_{ijt}) = f^M_t(X_{ijt})e^{-u^M_{it}}, \qquad \forall \text{ j, i, t} \tag{3}$$

where $u^M_{it} \geq 0$ and $f^M_t(\cdot) \geq f^j_t(\cdot)$ so that the metafrontier envelops the family of group frontiers closely from above. As for the group frontiers, the use of the stochastic frontier regression and the assumption of a truncated-normal distribution, $N(\mu^j(Z_{ijt}), \sigma_u^{j2}(Z_{ijt}))$, of the u_{ijt} term allows for the evaluation of factors that explain the technology gap ratio (Bravo-Ureta, Higgins, and Arslan 2020). But it should be noted that these are different factors than the ones affecting farm-level efficiency (Huang, Huang, and Liu 2014).

In the stochastic metafrontier model the TGRs are the efficiency scores generated when fitting estimating Equation (3) and these are obtained from Equation (4).

$$TGR^j_{it} = \frac{f^j_t(X_{ijt})}{f^M_t(X_{ijt})} = e^{-u^M_{it}} \leq 1 \tag{4}$$

Estimates of meta-technical efficiency (MTE) are extracted by calculating the product of the technology gap ratio (TGR) and group technical efficiencies (TE).

$$MTE^j_{it} = TGR^j_{it} * TE^j_{it} \tag{5}$$

Note that the decomposition in Equation (5) is different from the one proposed in Battese, Rao, and O'Donnell (2004) or any subsequent applications. In semi-parametric metafrontier models, stage 2 involves the original dependent variable in the frontier and the non-parametric measure of inefficiency obtained represents the full distance between the actual observation and the metafrontier. They isolate the residual gap between the group and metafrontiers and the smaller meta efficiency is divided by the somewhat larger group efficiency. The innovation in Huang, Huang, and Liu (2014) was to replace the dependent variable with estimated variables in the second stage regression. But here, in order to measure technology gaps directly and therefore retrieve a measure similar to the meta-technical efficiency defined in Battese, Rao, and O'Donnell (2004), the product of stage 1 and stage 2 efficiencies is used as each part of expression 5 lies between zero and one, $0 \leq TE^M \leq 1$.

The interpretation of the technology gap ratio is as difficult as it was previously. For example, a gap of say TGR = 0.80 means that a given frontier is within 20% of the metafrontier, and since gaps are farm-level observations, most studies report the average gap between groups and their metafrontiers. In a global analysis of country performances these gaps are likely to be much wider than in a local analysis where "global" best practice is locally defined. Group and meta efficiencies are also expected to vary inversely with sample size, with smaller groups resulting in higher general efficiency levels than larger groups.

2.2 Groups in this dataset

Metafrontier analysis begins with forming groups by homogenous technologies. In the literature geographic demarcations are commonly used (e.g., O'Donnell, Rao, and Battese 2008; Temoso, Hadley, and Villano 2015) and otherwise producers are classified by type (e.g., Gwebu and Matthews

2018). In this study $J = 4$ geographically defined groups were obtained from existing benchmarking studies.

The Karoo Management Panel contributed most of the data. This survey dataset for the Central Karoo District Municipality currently includes eight waves, for the period 2012–2019. Karoo farms specialise in small stock production and the sample includes wool and mutton flocks (Conradie and Landman 2015). Crop production is rare and generally limited in extent. Non-parametric benchmarking was applied to Wave 1 (Conradie and Piesse 2015), a stochastic frontier trend model was used for waves 1–3 (Conradie, Piesse, and Stephens 2019) and a Battese and Coelli (1995) stochastic frontier with technical inefficiency effects measured and explained inefficiency levels in waves 1–4 (Conradie 2019). In these studies, the local municipalities of Laingsburg and Beaufort West were pooled. There is not yet an analysis of the full eight waves of which the last four represent a serious drought. In waves 1–4 efficiency improvements were associated with the presence of wooled sheep, part-time farming and a diploma from *Grootfontein* Agricultural College. The courses for this diploma include modules in sheep and wool production and are tailored to the needs of Karoo farmers. The Karoo Management Panel covers a steep precipitation gradient with expected rainfall varying from 113 mm per annum (p.a.) at Laingsburg to 238 mm p.a. at Beaufort West. For this study the sample was split by district into 163 observations for Beaufort West and 248 observations for Laingsburg. Descriptive statistics are in Table 1.

The other two panel datasets are study group rather than survey data. These samples are expected to produce higher average levels of efficiency than the Beaufort West and Laingsburg groups because they are smaller and subject to self-selection with participants themselves choosing to participate in the study groups.

The Swellendam sample of 93 observations from ten farms over ten years is part of a rotating panel collected by the credit division of *Sentraal-Suid Kooperasie*. The Williston sample contains 75 observations collected from nine farms over thirteen years. Swellendam, in the Overberg, receives sufficient rainfall to support dryland crop production and, in this sample, wooled sheep is a secondary enterprise. Despite the small sample size Conradie and Genis (2020) estimated a Cobb Douglas stochastic frontier with inefficiency effects for the entire mixed farming system for these data. Only data on the sheep enterprise feature in this study and overhead costs were allocated to sheep according to its share of gross farm income. It was expected that Swellendam would set the standard because the levels of affluence often observed in the area facilitates the adoption of new technologies. In addition, sheep farming in this area is closely associated with pasture-based dairy farming, the only agricultural industry for which technical progress has been documented in South Africa in recent years (Vink, Conradie, and Matthews 2022).

The Williston sample includes nine farms and 75 observations for the period 2010–2021, and although this district is in the Karoo, due to selection bias the Williston group is expected to

Table 1. Descriptive statistics of farm-level inputs and output and selected other farm descriptors.

	Beaufort West $n = 163$	Laingsburg $n = 248$	Swellendam $n = 93$	Williston $n = 75$
Production function (R1000)				
Wool & mutton income	814^{a}	477^{b}	3121^{c}	1907^{d}
Stock sheep	1208^{a}	782^{b}	2568^{c}	$1086^{a,b}$
Feed & remedies	110^{a}	148^{a}	889^{b}	610^{c}
Fuel & machinery repairs	127^{a}	133^{a}	203^{b}	164^{a}
Estate repairs	72^{a}	38^{b}	$47^{a,b}$	$60^{a,b}$
Labour	315	170	196	250
Inefficiency sub-model				
Annual rainfall	153^{a}	136^{a}	418^{b}	134^{a}
D_ Grootfontein	4%	11%		68%
D_ part-time farming	9%	15%		29%

Notes: Financial figures in constant 2020 prices in R1000, with ZAR 17 = US$ 1 in August 2022.
Different superscripts indicate statistically significant differences across regions at $p < 0.05$. These were ANOVA tests.

outperform Laingsburg which has the same rainfall and perhaps even Beaufort West with almost twice the rainfall. The Williston farms specialise in mutton production and, as in Beaufort West and Laingsburg, crop production is uncommon. Since the area received less than 70% of expected annual rainfall since 2016, the Williston group adapted by setting feedlot systems for ewe-lamb production (Conradie and Geyer 2021). High feed costs forced these producers to cull aggressively, which should have increased efficiency, and strong vertical integration in the local cooperative of which these producers are members, brings cost and product price benefits.

Table 1 shows that the main differences between these four areas are flock size, and feed and animal health requirements. There was less variation between groups in fuel expenditure and mechanisation costs as well as in estate repairs and maintenance and there is no difference between groups in cost of labour. Mean farm income from sheep was different in each of the four groups, with the highest income recorded in Swellendam, followed by Williston and Beaufort West and Laingsburg in last place. The variation is partly explained by rainfall differences and partly by differences in farm and flock size. Based on these results it was hypothesised that the technology gap between the Laingsburg group frontier and the metafrontier might be the largest, the Beaufort West and Williston group frontiers slightly ahead of the Laingsburg frontier, with the Swellendam group frontier expected to be closest to the metafrontier.

2.3 Empirical models and specification tests

Group and metafrontiers were estimated using FRONT 4.1 (Coelli 1996) and the specification tests computed in Excel. The group frontiers' dependent variable was the natural logarithm of revenue from sheep and mutton in a given financial year. Nominal values were adjusted to constant 2020 prices with the sheep price index published in the Abstract of Agricultural Statistics (DALRRD 2021) and the undefined logged values of zero observations were replaced with zeros. Input price indices were from the same source. The main input is flock size, typically representing breeding ewes and replacement ewes in mutton flocks, and sometimes including whethers in the case of wooled sheep. Land is omitted as it is highly correlated with livestock numbers in pastoral systems. Land enhancing inputs are represented by a composite variable of expenditures on purchased and produced animal feed, remedies and reproductive costs such as artificial insemination (AI) and pregnancy checking. The model also includes repairs and maintenance of fixed farm infrastructure (*fences*) as well as labour and labour enhancing inputs. The latter is measured as the expenditure on fuel and machine repairs and maintenance. Theory requires positive coefficients on these inputs and in a translog specification the coefficients on pairs of complementary inputs will be positive while substitutes will have negative coefficients. The inefficiency model in use at the group stage simply included rainfall and two dummy variables for a specific type of training and part-time farming status. Rainfall and training were expected to increase efficiency while part-time farming was expected to decrease efficiency.

Hypothesis tests confirmed that the group models were indeed frontiers and to investigate if the Cobb Douglas restriction of the more general translog functional form was acceptable. These used general likelihood ratio tests where the test statistic $\lambda = -2 \times [L(H_0) - L(H_a)]$, and $L(H_0)$ and $L(H_a)$ are the log likelihood values of the unrestricted (H_0) and restricted (H_a) models, as described in Battese and Coelli (1988). Lambda is χ^2 distributed with degrees of freedom equal to the number of restrictions imposed.

The second stage of a SMF estimation requires that the pooling hypothesis must be rejected, and in this test the parameters obtained from a model estimated using the pooled data are treated as restrictions of a set of parameters obtained from fitting $J = 4$ separate models. In this study the number of restrictions was thus three times the total number of parameters (including intercepts and gamma) in the frontier sub-models in stage 1. In this test, model specification must be the same across all groups, and is therefore constrained by the performance of the group model with the smallest number of observations. Therefore, the Cobb Douglas restriction only needed to be investigated for this group. To replace actual output in Stage 2, estimated dependent variables $\hat{f}_t^j(X_{ijt})$ are the output for the i^{th} farmer in the t^{th} period for the j^{th} group based on the estimated coefficients of Equation (1) for each group. These estimates include the value of the intercept but omits statistical noise and any of the parameters of the inefficiency sub-model. Derived values for the J groups were pooled to estimate the meta frontier as specified in Equation (3) and its efficiency estimates were exported to compute the meta efficiencies. Rainfall was retained as an inefficiency variable in Stage 2. To capture institutional differences rainfall was supplemented by three district dummy variables with Laingsburg as the base-case, and the coefficients on these were expected to be negative because Laingsburg was one of the driest areas and appeared to show the least amount of innovation.

3. Results and discussion

3.1 Group frontiers

At the group stage, we modelled the best production functions for sheep farming that the meagre degrees of freedom allowed. Despite this limitation, there was clear evidence of inefficiency in all four groups. Gamma was significant and in fact very close to one in each case, suggesting that the observed farm-level inefficiencies were more than adequately represented by the three factors for which there were data across the four groups. In addition, the formal LR test of the joint significance of gamma and the coefficients of the Z-variables were easily rejected for Beaufort West, Laingsburg and Williston. This is shown in Table 2. The test statistic for Swellendam was smaller but since its inefficiency sub-model was simpler as well, the frontier hypothesis that the data is adequately represented by the mean response function was rejected for this group as well. The functional form for all the group frontiers was Cobb Douglas because the hypothesis that the coefficients on the squared and cross products introduced in the more general translog function were jointly zero in the case of Williston, and this determined the group frontier specification in all four cases.

The Beaufort West frontier model produced a reasonable fit with significant positive coefficients on three of the five inputs. In extensive livestock, production livestock is usually the most important factor of production, but in this case the coefficient on fuel and mechanisation was 68% larger than the coefficient on stock sheep. The labour input produced a commonly encountered coefficient of 0.05–0.10. The two inputs that failed to explain any of the variation in output were feed and remedies and estate repairs and maintenance. Excluding these coefficients, the sum of significant coefficients of 0.929 indicated mildly decreasing returns to scale, a typical finding in this production system (Conradie, Piesse, and Stephens 2019). In the inefficiency sub-model, the coefficients on all three Z-variables were significant, with all three factors contributing significantly to farm efficiency. The mean group efficiency for Beaufort West was 0.435 with 12% of the estimates above 0.800 and 23% below 0.200.

Table 2. Specification tests by group.

	Beaufort West	Laingsburg	Swellendam	Williston
Test 1				
Hypothesis	$\gamma = \delta_i = 0$	$\gamma = \delta_i = 0$	$\gamma = \delta_i = 0$	$\gamma = \delta_i = 0$
Lambda	252.22	378.37	9.73	22.28
Restrictions	5	5	3	5
Critical value ($p < 0.05$)	10.371	10.371	7.045	10.371
Test 2				
Hypothesis				$\beta_{ij} = 0 \ \forall\ i,\ j$
Lambda				16.97
Restrictions				15
Critical value ($p < 0.05$)				24.384

Three of the inputs had significant coefficients in the frontier model for Laingsburg, but output was not related to farm expenditure on fuel and mechanisation and hired labour. Smaller farms in the Laingsburg group, and extreme frugality during the drought, explain the small coefficient on fuel and mechanisation, while more use of family labour accounts for the lack of correlation between output and wages. In this case, part-time farming contributed to inefficiency. The largest input elasticity was on stock sheep, followed by feed, but at 0.72 the sum of the coefficients indicated even stronger decreasing returns to scale than for Beaufort West. The mean group efficiency of 0.497 varied between zero and 0.916 and there were 14% of observations each below 0.200 and above 0.800.

It is clear from Table 3 that other factors drive the success of sheep farming in Swellendam than mere livestock numbers, although efficiency was still positively correlated with rainfall. This could include more emphasis on optimal feeding to minimise inter-lamb periods and ewes' lifetime reproductive performance. In this model, labour was the dominant input followed by fuel and mechanisation and the sum of significant coefficients pf 0.577 indicated strongly decreasing returns to scale. In this group efficiency varied from 0.285 to 1.000 around a mean value of 0.512. There were only three observations above 0.800 and none below 0.200.

The model for Williston had an extraordinarily good fit for its small sample size. It is the only group frontier in which four of the five inputs were significant in the production function and the coefficients on these four variables suggest mildly increasing returns to scale of 1.069. The number of ewes in the flock was the most important input, followed by fuel and mechanisation and the labour input. The coefficient on feed and animal remedies of 0.087 had the highest value on this variable of the four groups, suggesting that feeding is a key feature of the Williston production system. The signs on the inefficiency terms (Z-variables) were as expected. At 0.532, mean efficiency in Williston was slightly higher than in Swellendam and the range was similar, varying from 0.139 to 1.000. There were two observations below 0.200 while almost 10% of the observations had efficiencies of >0.800, which suggests high levels of technical change in this group supporting increased efficiency.

3.2 Poolability

Pooling observations from all four groups produced a sample of 579 observations and estimating the common group frontier model generated a log likelihood value of −833.36. These results appear in the last two columns of Table 4 below. The only unexpected aspect of the results from this model was that the dummy variable for part-time farming produced a positive and significant coefficient, as in the Beaufort West model, but counter to the results for Williston and Laingsburg. The sum of the five input coefficients in this model indicated practically constant returns to scale.

The sum of log likelihood values of the four group models reported on in Table 3 was −598.62. This gave a test statistic on the LR pooling test of $\lambda = -2 \times [-833.36 - (-598.62)] = 508.72$. Since pooling imposed 31 restrictions on the system, the critical value for rejecting this hypothesis was $\chi^2_{0.05} = 43.773$. Thus, pooling was rejected and the metafrontier model was estimated in which the values from the group models replaced actual observations in the pooled sample.

3.3 The metafrontier model

The parameters from the metafrontier model are shown in the first two columns of Table 4. The large sample size is consistent with a good fit and the parameters obtained for the frontier variables were significant with the expected signs and relative magnitudes. However, they were much smaller than those obtained in the group model applied to the pooled data or in previous pooled models estimated on the Karoo Management Panel data (e.g., Conradie 2019). These results indicate an average estimate of returns to scale of just 0.48, which means that most of the sheep farms in this sample are substantially too large and/or too intensive. The inefficiency sub-model generated four significant coefficients, all with the expected signs. Although significant, the rainfall effect was smaller than anticipated and together with the unspecified institutional differences between

Table 3. Parameter estimates for the group stochastic frontier models.

	Beaufort West		Laingsburg		Swellendam		Williston	
	n = 163		n = 248		n = 93		n = 75	
	Coefficient	Std error	Coefficient	Std error	Coefficient	Std error	Coefficient	Std error
Production function								
Constant	5.324	0.917***	8.120	0.478***	8.439	1.901***	4.778	0.857***
ln(stock sheep)	0.326	0.033***	0.579	0.064***	0.000	0.038	0.547	0.134***
ln(feed & remedies)	−0.005	0.022	0.108	0.026***	0.012	0.020	0.087	0.042***
ln(fuel & mechanisation)	0.549	0.083***	0.018	0.046	0.163	0.085*	0.290	0.090***
ln(estate repairs)	−0.014	0.012	0.037	0.013***	0.005	0.031	0.017	0.017
ln(labour)	0.054	0.016***	0.005	0.023	0.414	0.126***	0.145	0.042***
Inefficiency model								
Constant	37.106	12.393***	−14.210	3.448***	2.671	0.645***	1.592	0.689***
ln(rainfall)	−18.098	5.848***	−6.588	2.159***	−0.329	0.121***	−0.139	0.150
D_Grootfontein training	−11.110	3.616**	−11.804	6.985*			−0.434	0.096***
D_parttime farming	−2.453	1.129***	13.871	6.255**			0.081	0.175
Variance and other model statistics								
Sigma squared	75.78	21.44***	52.63	13.34***	0.083	0.031***	0.149	0.022***
Gamma	9.99E-01	3.28E-04***	9.99E-01	5.07E-04***	1.00E + 00	6.93E-05***	1.00E + 00	2.11E-08***
Log likelihood statistic	−239.74		−325.65		−11.48		−21.75	

Note: *** 1% significance level, ** 5% significance level, * 10% significance level.

Table 4. Parameter estimates for the metafrontier and pooled models.

	Metafrontier model		Pooled group model	
	$n = 579$		$n = 579$	
	Coefficient	Std error	Coefficient	Std error
Production function				
Constant	11.164	0.243***	5.468	0.459***
ln(stock sheep)	0.318	0.023***	0.456	0.039***
ln(feed & remedies)	0.031	0.009***	0.077	0.012***
ln(fuel & mechanisation)	0.068	0.021***	0.367	0.047***
ln(estate repairs)	0.023	0.005***	0.017	0.011
ln(labour)	0.039	0.008***	0.043	0.015***
Metafrontier inefficiency				
Constant	1.645	0.145***		
ln(rainfall)	−0.065	0.028**		
D_Beaufort West	−0.561	0.044***		
D_Swellendam	−2.482	0.132***		
D_Williston	−0.928	0.052***		
Group inefficiency model				
Constant			10.033	3.347***
ln(rainfall)			−7.872	0.407***
D_Grootfontein training			−12.055	1.279***
D_parttime farming			5.646	1.288***
Variance and other model statistics				
Sigma squared	0.134	0.009***	36.90	4.06***
Gamma	0.174	0.032***	9.97E-01	6.33E-04***
Log likelihood statistic	−221.01		−833.36	

Note: *** 1% significance level, ** 5% significance level, * 10% significance level.

districts, gamma, indicated that inefficiency effects account for just 17.4% of the overall error variation.

3.4 Interpretation of the efficiency estimates from the metafrontier model

The purpose of Table 4 is to generate the performance data that have been reported in Table 5 and which are the key results of the study. Table 5 summarises overall (meta-) efficiencies per district and show how these meta-efficiencies are decomposed into within and between group performances. In each location, estimates for the best and worst case, as well as the average group member, are reported and these data allow a ranking of the districts relative to each other. Every previous benchmarking study of productivity in the Karoo have only generated data comparable to column 1, which measures how well the individual farm compares to the group given the existing technology. In this paper it is possible to produce a ranking of each district's best practice relative to that of other districts' best practice and relative to the best available technology in the industry.

Table 5. A summary of group efficiencies and technology gaps by district.

District		Group efficiency	Technology gap ratio	Meta technology efficiency
Beaufort West	Mean	0.444	0.474	0.212
	Minimum	0.000	0.364	0.000
	Maximum	0.925	0.617	0.503
Laingsburg	Mean	0.497	0.264	0.132
	Minimum	0.000	0.230	0.000
	Maximum	0.916	0.302	0.255
Swellendam	Mean	0.512	0.982	0.503
	Minimum	0.285	0.971	0.279
	Maximum	1.000	1.000	0.982
Williston	Mean	0.532	0.677	0.359
	Minimum	0.139	0.534	0.103
	Maximum	1.000	0.781	0.723

As anticipated from the group stochastic frontier model results, Swellendam was the best performing metafrontier sheep-producing district overall with a mean meta-efficiency of 0.503 with a range of 0.279-0.982. The TGR data shows that when setting aside individual group inefficiency, the maximum gap between the Swellendam frontier and the metafrontier is 3% while the average technology gap between Swellendam and the metafrontier is around 2%. The lower end of the meta-efficiency range reveals significant room for improvement within the group. This is confirmed by the modest mean group efficiency score of 0.512 reported for Swellendam. On the other hand, the maximum meta-efficiency value of 0.982 recorded for Swellendam indicates that Swellendam sets the standard in sheep production in the Western and Northern Cape, and perhaps in the country. Innovations that originate in the Overberg spread to the Karoo, but this is unfortunate since the two regions have such different natural resource endowments and thus it is unlikely there is much to be gained by followers imitating benefit best practices used by the leaders (Conradie, Piesse, and Thirtle 2009).

Williston is by far the most efficient of the groups in the Karoo. The mean meta-efficiency is 0.359, which is almost 70% higher than in Beaufort West, which has a mean meta-efficiency of 0.212, and 172% better than mean meta-efficiency in Laingsburg. The difference in performance within groups is similar in all three Karoo districts and in terms of within group performance Williston ranks 7% ahead of Laingsburg and Laingsburg 12% ahead of Beaufort West. Thus, it is not particularly useful to benchmark simply locally as the real problem is the huge degree of separation between what the best farmers in Laingsburg are doing compared to the best farmers in Williston.

The technological gap between Williston and the metafrontier is 32%, which places Williston on average 30% ahead of Beaufort West whose mean TGR indicates a 53% gap between that group and the metafrontier. As suggested above, the technology in Laingsburg was found to be furthest behind the frontier, on average more than 70%. This implies that with a given input bundle, sheep farmers in Laingsburg produce on average only 26% of the output achieved by the best sheep farmers in the country. There is clearly an urgent need for the Laingsburg agricultural association or the local cooperative, *Koup Produsente Kooperasie,* to partner with the public extension service to improve productivity in the area. If these two producer organisations fail to improve the quality of the local production technology, individual producers will fail and with them the local organisations and eventually the village will not be economically viable. This can happen extremely rapidly, especially during a drought (Conradie, Piesse, and Stephens 2019). The leadership of *Williston Vleis Kooperasie* must be commended for the initiative they have shown during the past few years of extreme drought leading its members to innovate. It is imperative that other Karoo districts follow the example of best practice areas and build resilience to climate change in their own communities.

This paper has shown the importance of allowing for sources of variation in group-specific TGR like environmental variables beyond the control of farms in comparative studies between districts. Ng'ombe (2017) considered two plausible explanations for heterogeneous technological diversity in Zambia. The first was a rainfall gradient that increases in a north-easterly direction and the second were technological spill-ins from neighbouring countries that increase in a south-westerly direction. Results showed spill-ins to be more important than rainfall. In that paper, in Southern Province and Lusaka Province, smallholders are better connected to regional supply chains and infrastructure but have the lowest rainfall. However, these farms operated within 11–12% of the metafrontier. On the other hand, Northern and Luapula Provinces, which are remote but receive higher rainfall, are at a distance of between 83% and 92% from the national level metafrontier, due to the scarcity of modern inputs in these regions. In the present study, while rainfall was significant and explains the high levels of technical performance in Swellendam, when this is combined with the overall affluence of the region and the influence this has on technology adoption, it is clear that innovation is the most important factor increasing efficiency in this farming community. However, at the same time, rainfall does not explain the larger performance difference between Williston and Laingsburg as precipitation levels are the same.

3.5 Limitations

There are some limitations to this study. The most important are based on the randomness of the districts included and the size of some of the groups, which are a function of missing data. However, given these constraints, the results are surprisingly informative and it would be helpful to test the robustness of the results with more data. Since the study was restricted to only the western half of the country, it is not possible to make any general inferences about how the Karoo compares to the rest of the country and that what technological spillovers might exist between regions. A broader database compiled from nationally representative surveys would be hugely helpful. Another concern raised by this analysis is that existing geographical boundaries may be a poor proxy for real underlying differences that could be revealed for example by cluster analysis.

4. Conclusion

This study investigates the added value of stochastic metafrontier analysis compared with previous group frontier models. This is a relatively new approach and this paper is only the fifth metafrontier paper using regional data and the third focussing on South African data. The Karoo has been the subject of benchmarking studies before, but farm-level performance has never been compared with that of sheep farming in other regions. The technology gaps between the Karoo and the rest of the industry and even within the Karoo recommend that these insular communities look further afield if they are to benefit from spillovers in innovation elsewhere.

Of course, there are many applications that could be done replicating this approach. Individual studies of key agricultural industries would provide insights into which institutional factors create pockets of local excellence to determine the best way to allow best practices to migrate from innovators to other regions. This used to be the purview of the public extension service although in recent years this has lapsed or ceased altogether. If the public extension service fails to do so, the private sector should be encouraged to take over the role, perhaps following the example of Trace and Save in the dairy industry (Conradie, Galloway, and Renner 2022; Galloway et al. 2018).

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Nicolette Matthews http://orcid.org/0000-0002-8975-9026
Beatrice Conradie http://orcid.org/0000-0002-4675-6300
Jenifer Piesse http://orcid.org/0000-0002-6584-6678

References

Battese, G.E., and T.J. Coelli. 1988. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* 38, no. 3: 387–99.

Battese, G.E., and T.J. Coelli. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20, no. 2: 325–32.

Battese, G.E., D.S.P. Rao, and C.J. O'Donnell. 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.

Bravo-Ureta, B.E., D. Higgins, and A. Arslan. 2020. Irrigation infrastructure and farm productivity in the Philippines: A stochastic meta-frontier analysis. *World Development* 135: 105073.

Coelli, T.J. 1996. A guide to Frontier version 4.1: A computer program for stochastic frontier production and cost function estimation. CEPA working paper 7/96, Department of Econometrics, University of New England, Armidale, Australia.

Conradie, B. 2019. Designing successful land reform for the extensive grazing sector. *South African Journal of Agricultural Extension* 47, no. 2: 1–12.

Conradie, B., and A. Genis. 2020. Efficiency of a mixed farming system in a marginal winter rainfall area of the Overberg, South Africa, with implications for thinking about sustainability. *Agrekon* 59, no. 4: 387–400.

Conradie, B., and A. Geyer. 2021. *Williston voerstelsel: min ooie met baie lammers* [Williston feedlots: Few ewes with lots of lambs]. CSSR working paper 465, University of Cape Town.

Conradie, B., and A. Landman. 2015a. Wool versus mutton in extensive grazing areas. *South African Journal of Agricultural Extension* 43, no. 1: 22–31.

Conradie, B., and J. Piesse. 2015b. Productivity benchmarking of free-range sheep operations: Technical efficiency, correlates of productivity and dominant technology variants for Laingsburg. *South Africa. Agrekon* 54, no. 2: 1–17.

Conradie, B., C. Galloway, and A. Renner. 2022. Private extension delivers productivity growth in pasture-based dairy farming in the Eastern Cape, 2012–2018. *Agrekon* 61, no. 2: 109–20.

Conradie, B., J. Piesse, and C. Thirtle. 2009. District-level total factor productivity in agriculture: Western Cape province, South Africa, 1952–2002. *Agricultural Economics* 40, no. 3: 265–80.

Conradie, B., J. Piesse, and J. Stephens. 2019. The changing environment: Efficiency, vulnerability and changes in land use in the South African Karoo, 2012–2014. *Environmental Development* 32: 100453.

Department of Agriculture, Land Reform and Rural Development (DALRRD). 2021. Abstract of Agricultural Statistics, Pretoria.

Galloway, C., B. Conradie, H. Prozesky, and K. Esler. 2018. Opportunities to improve sustainability on commercial pasture-based dairy farms by assessing environmental impact. *Agricultural Systems* 166: 1–9.

Gwebu, J.Z., and N. Matthews. 2018. Metafrontier analysis of commercial and smallholder tomato production: A South African case. *South African Journal of Science* 114, no. 7-8: 55–62.

Hayami, Y., and V. Ruttan. 1971. *Agricultural development: An international perspective*. Baltimore: Johns Hopkins University Press.

Huang, C.J., T.H. Huang, and N.H. Liu. 2014. A new approach to estimating the metafrontier production function based on a stochastic frontier framework. *Journal of Productivity Analysis* 42: 241–54.

Kawagoe, T., Y. Hayami, and V. Ruttan. 1985. The intercountry agricultural production function and productivity differences among countries. *Journal of Development Economics* 19, no. 1-2: 113–32.

Ng'ombe, J.N. 2017. Technical efficiency of smallholder maize production in Zambia: a stochastic meta-frontier approach. *Agrekon* 56, no. 4: 347–65.

Nyam, Y.S., N. Matthews, and Y.T. Bahta. 2020. Improving livelihoods of smallholder farmers through region specific strategies: a case study of South African sheep production. *Agrekon* 59, no. 1: 1–15.

O'Donnell, C.J., D.S.P. Rao, and G.E. Battese. 2008. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics* 34: 231–55.

Temoso, O., D. Hadley, and R. Villano. 2015. Performance measurement of extensive beef cattle farms in Botswana. *Agrekon* 54, no. 4: 87–112.

Vink, N., B. Conradie, and N. Matthews. 2022. The economics of agricultural productivity in South Africa. *Annual Review of Resource Economics* 14: 131–149.