The Australian Journal of
Agricultural and Resource Economics

Journal of the Australian Agricultural and Resource Economics Society

The Australian Journal of Agricultural and Resource Economics, 55, pp. 90–106

Event-specific data envelopment models and efficiency analysis*

Robert G. Chambers, Atakelty Hailu and John Quiggin†

Most, if not all, production technologies are stochastic. This article demonstrates how data envelopment analysis (DEA) methods can be adapted to accommodate stochastic elements in a state-contingent setting. Specifically, we show how observations on a random input, not under the control of the producer and not known at the time that variable input decisions are made, can be used to partition the state space in a fashion that permits DEA models to approximate an event-specific production technology. The approach proposed in this article uses observed data on random inputs and is easy to implement. After developing the event-specific DEA representation, we apply it to a data set for Western Australian barley production data. Our results highlight the need for acknowledging stochastic elements in efficiency analysis.

Key words: DEA, efficiency, event-specific DEA, event-specific technology, risk, stochastic technology.

1. Introduction

Agricultural production technologies are inherently uncertain. Unpredictable climatic variables such as rainfall are essential to production, and farmers must plan for a range of contingencies when making production decisions.

With very few exceptions, models designed for making efficiency comparisons model deterministic technologies, and stochastic influences are confined to an error term. O'Donnell *et al.* (2006) have shown that efficiency analysis, whether based on stochastic frontier (SFA) or data envelopment models, can be seriously biased if methods developed for nonstochastic technologies are applied to data sets generated by firms facing truly stochastic technologies and decision environments.

This article demonstrates how data envelopment analysis (DEA) methods can be adapted to accommodate truly stochastic decision environments, such as those encountered by farmers, in a state-contingent setting. Specifically, we

* The authors would like to thank Rodney Hunter and Ross Kingwell (Department of Agriculture Western Australia) for kindly providing the agricultural experiment data used in this article.

† Robert G. Chambers is at the Department of Agricultural and Resource Economics, University of Maryland, Maryland, USA. Atakelty Hailu (e-mail: atakelty.hailu@uwa.edu.au) is at the School of Agricultural and Resource Economics, University of Western Australia, Crawley (Perth), Western Australia, Australia. John Quiggin is at the School of Economics, University of Queensland, Queensland, Australia.

doi: 10.1111/j.1467-8489.2010.00517.x

show how observations on a random input, such as rainfall or other climatic variables, can be incorporated in DEA specifications in a fashion that recognises the essential stochastic nature of agricultural production.

The central idea is to define the state space describing the stochastic decision environment faced by farmers in terms of a random input. Once the state space is defined, observations on the random input are used to partition the state space in a fashion that permits DEA models to approximate an event-specific production technology.

After developing the event-specific DEA representation, we apply it to a data set from Western Australia to illustrate the differences in efficiency calculations that can emerge when the truly stochastic nature of agricultural production is taken into account. For our data set, the importance of accounting for the stochastic nature of agricultural production is highlighted by the dramatic consequences it has for calculated efficiency scores.

In what follows, we first define a stochastic production technology. Then, we show how information on a random input can be used to define a partition of the state space that permits specification of an event-specific version of the technology, and we show how that specification can be implemented in a DEA framework. We discuss our data set next, and then we apply our method to that data set and discuss our findings. This is followed by a discussion of results from a sensitivity analysis where the DEA efficiency estimates are bootstrapped and the results compared to those obtained from the standard DEA scores. Then, the article concludes.

2. The stochastic technology

The stochastic setting is formally represented by a measurable space (S, Ω) where S is the state space and Ω is its measurable subset (events). In this setting, random variables are treated as measurable maps from S to the reals. Thus, any random variable, $\tilde{f}$, can be thought of as the element of $\mathbb{R}^{S}$, defined by

$$\tilde{f} = \{f(s) : s \in S\},$$

where $f : S \rightarrow \mathbb{R}$ is the map defining the random variable, and it is required that $\{s : f(s) = \phi\}$ belongs to Ω for all $v \in \mathbb{R}$. Random variables will always be distinguished from their *ex post* (realised) values by a tilde (˜). Hence, $\tilde{f} \in \mathbb{R}^{s}$ represents the random variable, and f(s) denotes the *ex post* (observed) outcome associated with realisation s of S. Denote by $\tilde{1}$, the degenerate (constant) random variable whose outcome equals one for all $s \in S$.

The stochastic production technology uses multiple nonstochastic inputs to produce a single stochastic output.[1] That stochastic output is represented by

[1] We concentrate on a single output technology for the sake of simplicity. It is apparent, however, that our method can be easily extended to multiple output stochastic technologies following the lines developed in Chambers and Quiggin (2000).

the random variable $\tilde{z} \in \mathbb{R}^S_+$. The technology is represented by a set $T \subset \mathbb{R}^S_+ \times \mathbb{R}^N_+$, where N represents the number of inputs that are under the direct control of the producer and that are applied prior to the resolution of uncertainty. T is defined by

$$T = \{(\tilde{z}, x) : x \text{ can produce } \tilde{z}\},$$

where $\tilde{z} \in \mathbb{R}^S_+$ denotes the stochastic output and $x \in \mathbb{R}^N_+$ denotes the nonstochastic inputs under the control of the producer. We assume that T is non-empty, exhibits free disposal of inputs and outputs and is convex. The interpretation of the technology is as follows. Before the producer knows the realisation $s \in S$, he or she picks ($\tilde{z}$, **x**) from within T. After the producer makes this choice, then a neutral player, 'Nature', makes a choice from S. If the realised state is $s \in S$, then realised output is $z\,(s)$; while if $s' \neq s$ is realised, then *ex post* output is $z\,(s')$.

We now consider a comprehensive partition of the state space, S, into mutually exclusive events. Call that partition $\hat{\Omega}$ and denote a typical element of it by ω. These events are mutually exclusive

$$\omega \neq \omega' \Rightarrow \omega \cap \omega' = \varnothing$$

and the partition is comprehensive

$$\bigcup_{\omega \in \hat{\Omega}} = S.$$

If one only has data on *ex post* output realisations, then empirical approximation of T in a practical data setting requires an identifying restriction on T. To that end, we assume that T can be represented in terms of a family of event-specific stochastic production functions so that

$$T = \{(\tilde{z}, x) : z(s) \leq g_\omega(x, s), \omega \in \hat{\Omega}, s \in \omega\},$$

where each g_ω is a nondecreasing and concave function of the nonstochastic inputs. In what follows, g_ω is termed the *event-specific production function* for the event ω. The basic idea behind an event-specific representation of the technology is that the occurrence of different events fundamentally changes the *ex post* conditions under which stochastic production takes place. An obvious special case of an event-specific production function is the state-contingent production function that has been axiomatically studied by Chambers and Quiggin (2000). In that case, $\hat{\Omega} = S$ and

$$T = \{(\tilde{z}, x) : z(s) \leq g_s(x, s), s \in S\}.$$

An event-specific technology has a number of advantages for applied work. Most importantly, as noted earlier and as we show in the following, it allows one to use *ex post* observations on output in the construction of empirical approximations of the technology. Thus, choosing an event-specific representation represents an important identifying restriction. However, it also comes with significant costs. Perhaps most importantly, it makes a very strong assumption upon how producers can react to uncertainty. As pointed out by Chambers and Quiggin (2000), it requires that inputs cannot be allocated differentially to prepare for different stochastic outcome. Moreover, as O'Donnell, Chambers and Quiggin (2010) have shown through simulation analysis, if the true technology is not event-specific, then empirical representations of the technology based upon this identifying restriction can lead to serious errors and biases in approximating the frontier of the technology and in measuring efficiency. Theoretically, as Chambers and Quiggin (2000) have demonstrated, event-specific technologies place strong a priori assumptions on the degree of substitutability between *ex post* realisations of the stochastic output.

The special case of the event-specific technology, known as the state-contingent production function, is the most common empirical representation of stochastic technologies. It forms the basis for the standard representation of most SFA representation of technologies. As a general rule in applied econometric work, however, the practical specification of S is predicated more upon econometric and empirical convenience than upon capturing the actual decision environment that the decision-maker faces. More specifically, S is usually viewed as an 'error' space that arises from problems in measuring inputs and outputs and simple, although econometrically convenient, stochastic errors by producers who face a nonstochastic decision environment.

But in the truly stochastic decision environment in which most firms operate, S is not an 'error space'. Rather, S provides a comprehensive and mutually exclusive description of *all possible states of the world* that the producer can face after he or she makes his or her decision about the nonstochastic inputs $\mathbf{x}$ and the stochastic output $\tilde{z}$. In many practical instances, S can be relatively narrowly defined. For example, for many agricultural technologies, the main stochastic factors affecting realised output are climatic in nature, be it adequate moisture, absence of frost and early freeze, or other random inputs to the production process usually captured in our notion of weather. In those instances, the elements of S would be associated, for example, with the different levels of future rainfall that are practically possible when the crop is planted. And 'events' would be associated with different possible ranges of rainfall, say, low, medium or high. In each of these cases, it is important to realise that the *ex post* realisations of these random inputs are chosen by Nature and not by the producer. Thus, while they are inputs in one sense, they are not inputs chosen by the producer and thus should not be included in $\mathbf{x}$, which are the inputs that the producer has under his or her control. Their effect on physical outcomes enters through the state-contingent nature of the technology as a result of a separate choice by the neutral actor, Nature, and not as a result of producer choice.[2] Therefore, in what follows, we assume that S can be defined by the possible realisations of a real-valued, random input, which with an abuse of notation we denote as s, to the production process whose realisation occurs after ($\tilde{z}$, $\mathbf{x}$) is chosen. Hence, in what follows, $S \subset \mathbb{R}_+$ corresponds to the support of that random input. The partition of S given by $\hat{\Omega}$ is then given by consecutive subintervals of the positive reals.

The choice of $\hat{\Omega}$ is motivated by the need to represent production uncertainty in a relatively compact and empirically tractable fashion. For practical purposes, this requires that the number of elements of the partition $\hat{\Omega}$ should be small. It does not mean, however, that our method can only be applied if there is only one random input to the production process. Suppose that there were two. Then, S could be defined as a subset of $\mathbb{R}^2_+$, and events could be defined by appropriate partitions of that set.

An alternative approach, as in Henderson and Kingwell (2005), is to treat rainfall as an input to production, estimate a technology incorporating a response function and apply DEA. This approach has the advantage that all the information in the rainfall data can be used, without the problem that, as partitions become finer, more firms are classed as efficient.

There are, however, some notable disadvantages. First, because rainfall levels are determined exogenously by Nature, realised *ex post* standard concepts of allocative efficiency are not applicable. Second, notions such as constant returns to scale are problematic. Third, the approach relies critically on the correctness of the estimated functional form. This is particularly problematic in the presence of von Liebig effects, resulting in a negative marginal product at high levels of rainfall. Such effects may also be present for endogenously chosen inputs such as fertiliser. In this case, however, they are less problematic, as farmers would never consciously choose to apply fertiliser at levels such that the marginal product is negative.

3. A DEA model of the event-specific technology

Our theoretical model relates *ex post* output to realisations of the random input, s, according to

$$z(s) \leq g_{\omega}(x, s),$$

$\omega \in \hat{\Omega}$, $s \in \omega$. Thus, in terms of a DEA technology, one can legitimately think in terms of a technology that characterises the interaction between nonstochastic inputs, the stochastic inputs, and realised output. A standard Variable Returns to Scale (VRS) DEA representation of such a technology, assuming free disposability, would be:

[2] Quiggin *et al.* (2010) discuss in more detail the relationship between discrete stochastic programming and state-contingent production.

$$T^D = \left\{ (z, x, s) : z \le \sum_{k=1}^{K} \lambda_k z^k, x \ge \sum_{k=1}^{K} \lambda_k x^k, \right.$$
$$s \ge \sum_{k=1}^{K} \lambda_k s^k, x \ge \sum_{k=1}^{K} \lambda_k = 1,$$
$$\left. \lambda_k \ge 0, k = 1, \ldots, K \right\}$$

where (z^k, $\mathbf{x}^k$, s^k) corresponds to the kth observation's *ex post* output, the nonstochastic inputs and the observed random input, and where $\lambda = (\lambda_1, \cdots \lambda_K)$ are the DEA activity variables.

Notice that while T^D accounts for the presence of the random input, it is not a proper event-specific technology because it presumes that the same production frontier applies across all events $\omega \in \hat{\Omega}$. A technology that accommodates both the presence of the random input, and the event-specific nature of the technology can be constructed by using the K *ex post* values of the random input to partition the data into subsets that correspond to each of the events ω defined by the partition $\hat{\Omega}$ of the state space. Denote the number of observations falling into the event ω by $K(\omega)$ and the kth observation falling into that event by ($z^k(\omega)$, $\mathbf{x}^k(\omega)$, $s^k(\omega)$). Then, the event-specific DEA frontier associated with those observations and with event ω is given by

$$T^\omega = \{ (z, x, s) : z \le \sum_{k=1}^{K(\omega)} \lambda_k z^k(\omega), x \ge \sum_{k=1}^{K(\omega)} \lambda_k x^k(\omega),$$
$$s \ge \sum_{k=1}^{K(\omega)} \lambda_k s^k(\omega), \sum_{k=1}^{K(\omega)} \lambda_k = 1,$$
$$\lambda_k \ge 0, k = 1, \ldots, K(\omega) \}$$

and the (VRS) DEA approximation to T is given by

$$T^{\hat{\Omega}} = \{ (z, x, s) : (z, x, s) \in T^\omega, \omega \in \hat{\Omega} \}$$

4. Data and application

To illustrate how a DEA approximation to an event-specific technology can be constructed and the difference that it can make in actual efficiency calculations, we apply our methodology to a data set on crop yields from experimental field trials. The data are described in the following and some summary statistics presented. In the subsequent subsection, we explain how the data is partitioned and discuss the results.

4.1. Data

We use a data set of 260 observations from field trials of a Stirling Barley variety conducted in the Wheatbelt region of Western Australia between 1991 and 1995. These field trials are part of the crop variety testing (CVT) program of the Department of Agriculture (Hunter 2005). The focus of the CVT trials changes with the stage of development in a crop line. The data used here are from the final stage or prerelease trials. The final stage trials have an agronomic focus and are aimed at evaluating the effects on yield and grain quality of practices such as fertiliser combinations and rates. Thus, the treatments in these experiments are designed to be close to actual farm practices. The data relate yields on barley production to three fertiliser inputs (nitrogen, phosphorus and sulphur) that were under the direct control of the experimenter and two inputs (pre- and postsowing rainfall) that were not under the control of the experimenters at the time that fertiliser applications were made. Thus, for the purposes of our analysis, we take the random inputs, s, defining the state space to be rainfall as measured in millimetres.

Summary statistics from our data set are presented in Table 1. The mean and median grain yield values in our sample are similar to the average yield for the region. These averages are also close to yield values in a sample of 65 actual farms that the authors used for comparison purposes. The range and degree of variability in phosphorus rates are also similar in both the experimental and actual farm data sets. However, the nitrogen rates in the experimental data have a lower mean but are more variable. This is partly due the fact that the actual farm data reports nitrogen application rates across all cereals (wheat, barley and oats) and is thus likely to indicate higher nitrogen application rates because nitrogen rates on wheat tend to be higher pulling the cereal averages up relative to barley nitrogen rates. Given these similarities in yield and phosphorus rates, and also given the fact that the nitrogen rates in the experimental data cover a wide range, there is no reason to suspect that the shape of the yield function in our experimental data are different from what one would expect in actual farm data.

There are several reasons why agricultural field trial data provide a particularly convenient framework in which to illustrate our methodology. First, because these data emerge from experiments by professional agronomists who are presumably well acquainted with the most modern and advanced production methods, it is hard to imagine that there should be any inherent technical efficiency (TE) differences across observations, other than those that

Table 1 Data summary statistics

Summary statistic	Median	Mean	Minimum	Max
Annual rain (mm)	326.00	375.10	105.50	1014.30
Presowing rain (mm)	90.60	104.12	0.00	325.00
Postsowing rain (mm)	256.10	270.90	30.00	784.20
Grain yield (kg/ha)	2191.37	2247.88	1.67	5162.06
Nitrogen (kg/ha)	18.90	16.52	0.00	69.05
Phosphorus (kg/ha)	10.50	10.39	0.00	29.23
Sulphur (kg/ha)	16.80	17.72	0.00	362.50

emerge from truly random effects and observation error. Thus, in principle, one would expect most such observations to be relatively close to the ideal frontier. This is not the case, for example, in data that are gathered under less controlled circumstances, where true differences in ability and in 'human capital' can explain observed efficiency differences. Second, these data contain inputs that are both under the direct control of the experimenters (fertiliser levels) and inputs that are controlled by Nature. Hence, they seem to offer an ideal framework in which to investigate how apparent efficiency differences can emerge across observations not from any inherent differences in knowledge or true efficiency but from the truly stochastic nature of such technologies.

4.2. Application

For the empirical application, we have partitioned the data using rainfall levels. Figure 1 presents the empirical distribution for rainfall over the observations in the sample. On the basis of this empirical distribution, we have split the rainfall state space into three events: *low rainfall* (below 277.2 mm per annum), *medium rainfall* (between 277.2 and 426.8 mm per annum) and *high rainfall* (above 426.8 mm per annum). These three groups have, respectively, 82, 91 and 97 observations in them.

As a reviewer correctly points out, this partitioning of our data set according to three rainfall events is to some extent *ad hoc*. One could legitimately argue in favour of partitioning the data into, say, four events or an even a higher number. More generally, one would expect that in most practical settings, the partitioning of the data sets and its effects on the ultimate results

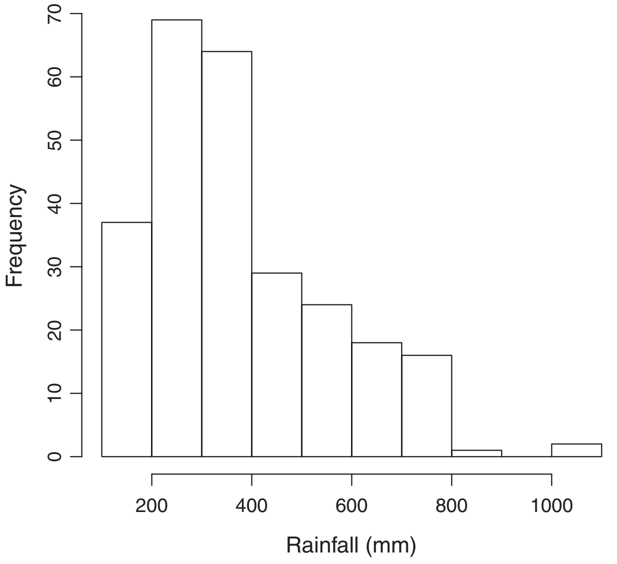

Figure 1 Empirical distribution of rainfall.

from the analysis should be investigated thoroughly by appropriate sensitivity analysis.

In making this partition, we intentionally attempted to choose the data partition so that there are roughly equal numbers of observations in each event. As studied by Zhang and Bartels (1998) and Fraser and Graham (2005), calculation of efficiency scores by DEA methods can be strongly affected by differences in sample size. For example, it is well recognised in the DEA literature that as the number of firms or decision making units decreases, the efficiency scores of the remaining units tend to rise. In the limit, therefore, one would expect that if the event partition were fine enough, virtually all firms would be declared efficient in an event-specific model.

In the empirical analysis, input-oriented (TEx) and output-oriented (TEy) technical efficiency scores were computed for the observations. First, we calculated the efficiency scores using representation T^D above that presumed that all observations come from a common technology (i.e. a combined frontier). Then, we calculated efficiency scores from frontiers calculated from the data partitioned according to the three rainfall events (low, medium, high) using the representation T^{ω}. In all cases, we presumed that the technology exhibited variable returns to scale and free disposability of inputs and outputs. Summary of these estimates are presented in Table 2.

We then compared the resulting efficiency scores that emerged from these two distinct methods using two different test statistics: the Kolmogorov–Smirnov nonparametric test and Banker tests. The Kolmogorov–Smirnov test statistic is a general distribution-free nonparametric test which quantifies differences in both location and shape of empirical cumulative distribution functions. Banker's test (Banker 1993), on the other hand, uses F-statistics that can be constructed from the TE estimates under the assumption of normal or exponential distributions for the efficiency terms (Banker 1993; Banker and Chang 1995), under the null hypothesis that T^D and $T^{\hat{\Omega}}$ are the same.[3]

According to the Kolmogorov–Smirnov test results (Table 3), the input- and output-oriented efficiency scores calculated relative to T^{ω} for ω equal to high rainfall are significantly higher (at 99% confidence level) than those calculated relative to T^D. Similar results are obtained for ω equal to medium rainfall. The Banker test results reported in Table 4 confirm these findings except in the case of the output-oriented scores for the medium rainfall group.

The input- and output-oriented efficiency scores calculated relative to T^{ω} for ω equal to low rainfall, however, are found to be similar to those calculated for these observations from T^D. Thus, on the basis of these results, we are led to conclude that T^D does a relatively good job of capturing the sto-

[3] These tests are based on analytical results obtained for the single output case. For multiple output efficiency scores, such analytically based statistical tests are not available and one has to rely on bootstrapping methods (Simar and Wilson 2000b).

Table 2 A Comparison of technical efficiency (TE) estimates from separate and combined technology frontiers

	Low rainfall group		Medium rainfall group		High rainfall group	
	Combined frontier	Separate frontier	Combined frontier	Separate frontier	Combined frontier	Separate frontier
Input-oriented TE						
Mean	0.88	0.88	0.72	0.89	0.44	0.90
Median	0.90	0.90	0.68	0.92	0.40	0.96
Output-oriented TE						
Mean	0.75	0.76	0.63	0.72	0.53	0.66
Median	0.73	0.75	0.62	0.73	0.52	0.65

Table 3 Kolmogorov–Smirnov tests of technical efficiency (TE) estimates from separate and combined frontiers (P-values for two-sided and one-sided tests)

Null hypothesis	Low rainfall group	Medium rainfall group	High rainfall group
Input-oriented/VRS			
TEs from combined and separate frontiers are same	0.998	0.0000	0.0000
TEs from combined frontiers are not smaller	0.7372	0.0000	0.0000
Output-oriented/VRS			
TEs from combined and separate frontiers are same	0.998	0.0247	0.0077
TEs from combined frontiers are not smaller	0.7372	0.0123	0.0038

VRS, Variable Returns to Scale.

Table 4 Banker test difference in efficiency scores from separate and combined technology frontiers (Note: figures indicate distribution area beyond critical statistic value, i.e. $P(X > x)$)

	Lower rainfall group	Medium rainfall group	High rainfall group
Under normal distribution assumption for efficiency terms			
TEx	0.46	0.00	0.00
TEy	0.34	0.40	0.10
Under exponential distribution assumption for efficiency terms			
TEx	0.44	0.00	0.00
TEy	0.36	0.41	0.02

TEx, input-oriented technical efficiency; TEy, output-oriented technical efficiency.

chastic technology for low rainfall observations, but fails to capture the event-specific nature of the technology for increased levels of rainfall. A closer look at these numbers reveals some interesting patterns.

First, as the test results aforementioned indicate, for the medium and high rainfall groups, efficiency scores are higher when the DEA frontier includes only observations from the group. The magnitude and proportion of efficiency score changes are most pronounced for the high rainfall group. All the input-oriented efficiency scores and 78% of the output-oriented scores for the high rainfall observations are strictly higher when efficiency is calculated relative to T^{ω} rather than T^{D} (Table 5). The corresponding figures for the medium rainfall group are 90% and 71%. These changes are less frequent in the case of the low rainfall group but are virtually nil in magnitude as the figures in Table 2 show. These ratios of input-oriented efficiency scores from separate and combined frontiers (TEx ratios) are plotted against rainfall measurements in Figures 2 and 3.

Second, the frequency and level of disparity between efficiency scores is greater for the input-based scores than for output-oriented scores (for both medium and high rainfall groups). For the high rainfall group, the TE ratios of the input-oriented scores from separate and combined frontiers have a mean (and also median) value of 2.25; these mean and median values are lower (1.08 and 1.28, respectively) for the output-oriented scores. The pattern

Table 5 Efficiency change count: proportion of technical efficiency scores that are strictly higher for separate than for combined frontiers

	Lower rainfall group	Medium rainfall group	High rainfall group
TEx	25.61	90.11	100.00
TEy	51.22	71.43	78.16

TEx, input-oriented technical efficiency; TEy, output-oriented technical efficiency.

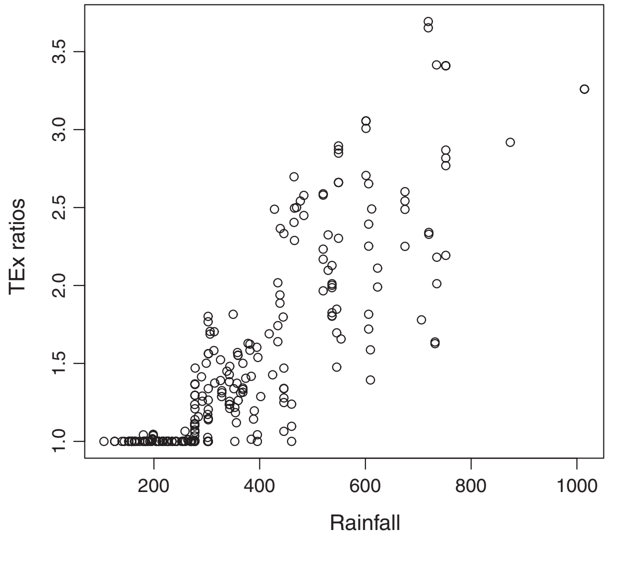

Figure 2 Ratios of input-oriented technical efficiency measures from separate and combined frontiers plotted against rainfall.

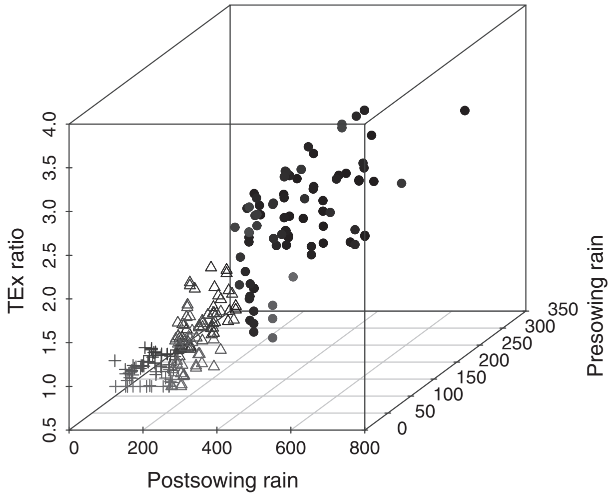

Figure 3 Ratios of input-oriented technical efficiency measures from separate and combined frontiers plotted against pre- and postsowing rainfall.

is the same for the medium rainfall group with the TE ratios from the input-oriented frontier being higher. However, the degree of efficiency understatement from T^D increases with rainfall in the case of input-oriented measures but not in the case of the output-oriented scores.

The observed pattern of efficiency underestimation associated with T^D can be explained as follows. When T^D is used, observations from the high rainfall category are dominated by those from the other two categories. In fact, for both input-oriented and output-oriented frontiers, none of the high rainfall observations are included as members of the best practice frontier for T^D. The pattern is less pronounced, but still observable, for the middle rainfall groups. This suggests that medium to high levels of rainfall fundamentally alter the production relationships between the inputs under producer control and rainfall variables. As we have noted earlier, we have imposed free disposability of inputs in the construction of DEA frontiers. However, it is very obvious that, in the extreme, very high levels of rainfall on a fixed plot of land can lead to a downward shift in productivity frontiers as the land becomes increasingly waterlogged. But even less dramatically, as rainfall reaches medium levels, there appears to be a levelling of the yield frontier associated with rainfall (in physical production terms, a von Liebig effect (Paris 1992; Chambers and Lichtenberg 1996)). Empirically, outliers in the data for rainfall levels that are low but not low enough to severely damage crop growth dominate observations from the high rainfall and medium rainfall groups that enjoy higher rainfall levels without correspondingly higher yield levels. Although rainfall levels in Western Australia are not very high, the yield plateau associated with von Liebig effects occurs within the range of the data. For soils with low water holding capacity, common in the West Australian Wheatbelt, additional rain mainly contributes to increased drainage (Asseng *et al*. 2001). A levelling of the yield frontier owing to a von Liebig type effect would naturally be associated with greater measured input inefficiency than measured output inefficiency.

5. Sensitivity evaluation through bootstrapping

The results reported so far rely on point estimates of efficiency scores that provide neither measures of the variability of the individual scores nor the strength of the evidence supporting the conclusions drawn regarding the impact of using event-specific formulations. To address these shortcomings, we use bootstrapping techniques. These techniques have been developed over the last decade (Simar and Wilson 2000a, 2008) and provide the most practical means of constructing confidence intervals around DEA scores.

The DEA bootstrap procedure used is the smoothed bootstrap described in Simar and Wilson (2008, 455–463). We generate bootstrapped samples for both the combined and event-specific frontiers. In each case, 5000 bootstrapped samples are used to generate a distribution of input-oriented and output-oriented efficiency scores for each observation. These samples are then used to construct confidence intervals and to undertake hypothesis tests of equality or differences in efficiency estimates from the different frontiers. The results are summarised in the following.

The 95% confidence intervals for the input-oriented efficiency scores are shown in Figure 4. In the figure, the observations are grouped by rainfall groups, and, within each group, observations are ordered by mean efficiency scores. It is clear from the figure that confidence intervals from the event-specific DEA model generally lie above the corresponding intervals from the combined DEA frontiers, with very little overlap in the case of high rainfall observations. There is also little overlap for most of the observations in the medium rainfall group. With the low rainfall group, on the other hand, the confidence intervals from the combined and event-specific DEA frontiers overlap almost everywhere. The results for the output-oriented measures are similar in pattern, with the least (most) overlap occurring for high (low) rainfall observations; see Figure 5.[4]

The evidence from these bootstrapped confidence intervals is consistent with the observations made above using the standard DEA scores: first, that efficiency scores from event-specific frontiers are higher than those obtained from the combined DEA frontier especially in the case of high and medium rainfall groups; and, second, that the effect of the frontier is more pronounced for the case of input-oriented rather output-oriented efficiency measures.

Finally, formal statistical tests on the bootstrapped samples of efficiency scores confirm the statistical test results reported in Tables 3 and 4 above.

[4] Another pattern observed in the plots is that the intervals are narrower for low efficiency observations than they are for high efficiency observations. This simply reflects the fact that scores for observations on or very close to the frontier are less certain.

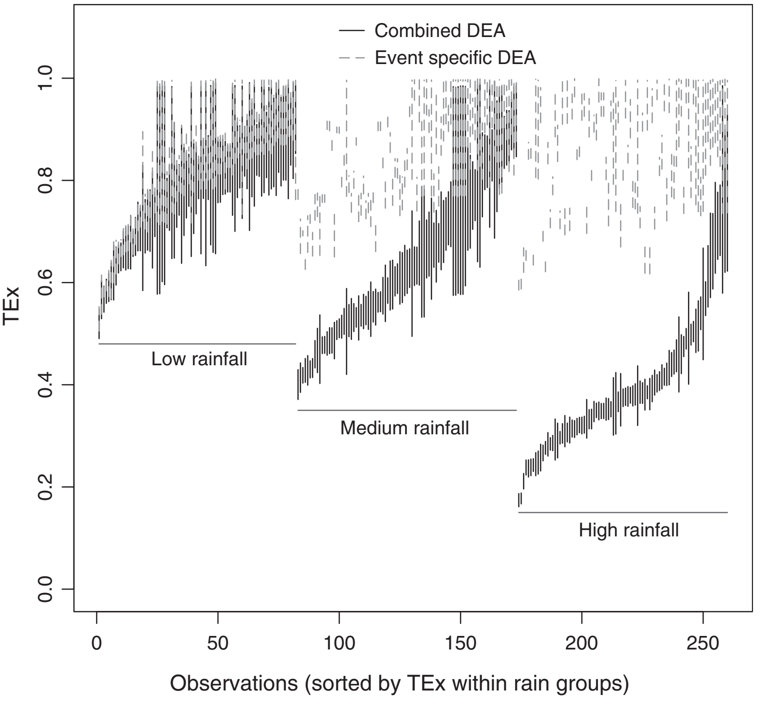

Figure 4 Bootstrapped confidence intervals for input-oriented efficiency estimates from combined and separate data envelopment analysis frontiers (α = 0.05, size of bootstrap = 5000).

Results from an observation-by-observation Kolmogorov–Smirnov hypotheses tests comparing the combined and separate DEA frontier scores are shown in Table 6. The hypothesis that the efficiency scores from the combined and event-specific DEA are equal is rejected by almost all the observations in our data set. This is the case for both input- and output-oriented efficiency measures. Further, the hypothesis that the combined efficiency score is not smaller than the score from the event-specific DEA frontier is rejected by most medium and high rainfall observations. These statistical test results confirm the test results reported in Table 3 based on the distribution of nonbootstrapped or standard DEA scores. These results are also consistent with the efficiency score change counts reported in Table 5. In summary, there is strong or statistically significant evidence that efficiency scores are underestimated when the event-specific nature of the production process is ignored.

6. Conclusion

When stochastic elements alter the nature of the underlying technology, efficiency measures computed from models that ignore these elements can lead to misguided management actions. The standard approaches to efficiency measurement do not allow for the stochastic nature of technologies. This is true of both deterministic approaches, such as DEA, and stochastic frontier

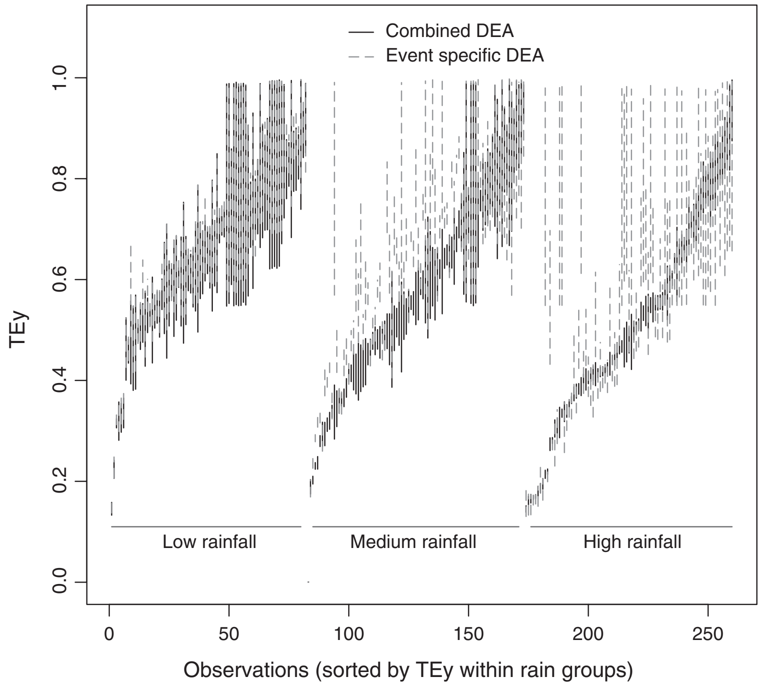

Figure 5 Bootstrapped confidence intervals for output-oriented efficiency estimates from combined and separate data envelopment analysis frontiers (α = 0.05, size of bootstrap = 5000).

(SFA) formulations. The latter incorporate stochastic errors merely as representations of measurement problems or omitted variables rather than as an explicit recognition of the stochastic nature of the underlying technology. Applying these models to data sets generated by a stochastic technology can lead to biased or erroneous estimates of efficiency performance. The purpose of this article is to show how event-specific representation of the production technology can be specified and then implemented within a DEA framework.

The article started by describing how the state space can be partitioned to define event-specific production relationships that approximate the underlying stochastic technology. The purpose of these event-specific technologies is to provide empirical representations of the underlying technology that reflect the fact that the structure of the production technology might be shaped differently by different events. The article then shows how the event-specific representations can be implemented in a DEA framework using the realised values of a random input to partition the data into comprehensive and exclusive subsets.

The event-specific DEA models are applied to agricultural field trial data, and the results compared with those obtained from a standard DEA model that ignores the stochastic nature of the data. These field trial data provide an excellent opportunity for demonstrating the benefits of the event-specific formulation. First, the trial data involve the use of inputs that are under the

Table 6 Proportion of observations for which hypothesis on technical efficiency (TE) scores are rejected: Results from the observation-by-observation Kolmogorov–Smirnov tests on bootstrapped efficiency estimates ($\alpha = 0.05$, size of bootstrap = 5000)

Null hypothesis	Low rainfall group (%)	Medium rainfall group (%)	High rainfall group (%)
Input-oriented/VRS			
TEs from combined and separate frontiers are same	100	100	100
TEs from combined frontiers are not smaller	9	88	94
Output-oriented/VRS			
TEs from combined and separate frontiers are same	100	99	100
TEs from combined frontiers are not smaller	49	84	70

VRS, Variable Returns to Scale.

direct control of the agronomist or the experimenter as well as inputs such as rainfall that are stochastic or under the control of Nature. Second, the experimental nature of the data implies that there is very little besides stochastic or natural events that would be responsible for observed efficiency differences. Rainfall data is used to partition the state space into low, medium and high rainfall events. Both input-oriented and output-oriented efficiency scores were calculated for the comparison of the alternative DEA models.

We find that estimates of efficiency performance change dramatically when an event-specific technology representation is adopted. This is particularly true for data points relating to medium and high rainfall events. For the data set used in the article, the calculations indicate that input-oriented efficiency scores were underestimated, on average, by 50% or more in the case of high rainfall event data. The results highlight the degree to which our understanding of efficiency levels can be distorted when models that do not recognise the stochastic nature of the production process are used.

The contribution in this paper may usefully be compared with the alternative approach of treating rainfall as a stochastic, and exogenously determined, input to a stochastic production function (Banker and Morey 1986), which has been applied to wheat production in Western Australia by Henderson and Kingwell (2005). For the case of an output-cubical technology, considered here, the two approaches are broadly equivalent. Unsurprisingly, therefore, the analysis here produces results broadly similar to those of Henderson and Kingwell (2005). When the state-contingent nature of production is taken into account, average efficiency scores increase, as does the number of firms classed as efficient. There are some differences between the approach adopted here and that of Henderson and Kingwell (2005). Henderson and Kingwell treat rainfall as a continuous variable and modify the standard efficiency analysis by treating the frontier for any given firm as being generated only by those firms with equal or lower rainfall. This approach offers a finer partition of the state space than that in the present study. On the other hand, the approach of Henderson and Kingwell incorporates the implicit assumption that the marginal product of rainfall is positive. The results of the present study suggest that von Liebig effects may be significant in the study area. It is therefore preferable to adopt a representation that allows for the high rainfall states to be either more or less favourable than lower rainfall states.

References

Asseng, S., Turner, N.C. and Keating, B.A. (2001). Analysis of water- and nitrogen-use efficiency of wheat in a mediterranean climate, *Plant and Soil* 233, 127–143.

Banker, R.D. (1993). Maximum likelihood, consistency and data envelopment analysis: a statistical foundation, *Management Science* 39, 1265–1273.

Banker, R.D. and Chang, H. (1995). A simulation study of hypothesis tests for differences in efficiencies, *International Journal of Production Economics* 39, 338–354.

Banker, R.D. and Morey, R.C. (1986). Efficiency analysis for exogenously fixed inputs and outputs, *Operations Research* 34, 513–521.

Chambers, R.G. and Lichtenberg, E. (1996). A nonparametric approach to the von liebig-paris technology, *American Journal of Agricultural Economics* 78, 373–386.

Chambers, R.G. and Quiggin, J. (2000). *Uncertainty, Production, Choice and Agency: The State-Contingent Approach*. Cambridge University Press, New York.

Fraser, I.M. and Graham, M. (2005). Efficiency measurement of Australian dairy farms: National and regional performance, *Australian Agribusiness Review* 13, 1–18.

Henderson, B. and Kingwell, R. (2005). Rainfall and farm efficiency measurement for broadacre agriculture in south-western Australia, *Australasian Agribusiness Review* 13, paper 14.

Hunter, R. (2005). *Crop Variety Testing Database: Background*. Department of Agriculture, South Perth, WA.

O'Donnell, C., Chambers, R.G. and Quiggin, J. (2006). *Efficiency Analysis in the Presence of Uncertainty*. Technical report. Risk and Sustainable Mangement Group, Schools of Economics and Political Science, University of Queensland, Brisbane.

O'Donnell, C.J., Chambers, R.G. and Quiggin, J. (2010). Efficiency Analysis in the Presence of Uncertainty, *Journal of Productivity Analysis* 33, 1–17.

Paris, Q. (1992). The von liebig hypothesis, *American Journal of Agricultural Economics* 74, 1019–1028.

Quiggin, J., Adamson, D., Chambers, S. and Schrobback, P. (2010). *Climate Change, Uncertainty and Adaptation: The Case of Irrigated Agriculture in the Murray-Darling Basin in Australia*. Technical report. Risk and Sustainable Mangement Group, Schools of Economics and Political Science, University of Queensland, Brisbane.

Simar, L. and Wilson, P.W. (2000a). A general procedure for bootstrapping in nonparametric frontier models, *Journal of Applied Statistics* 27, 779–802.

Simar, L. and Wilson, P.W. (2000b). Statistical inference in nonparametric frontier models: the state of the art, *Journal of Productivity Analysis* 13, 49–78.

Simar, L. and Wilson, P.W. (2008). Statistical inference in nonparametric frontier models: recent developments and perspectives, in Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds), *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, New York, pp. 421–521.

Zhang, Y. and Bartels, R. (1998). The effect of sample size on the mean efficiency of DEA with an application to electricity distribution in Australia, Sweden and New Zealand, *Journal of Productivity Analysis* 9, 187–204.