Measuring the Impacts of Production Risk on Technical Efficiency: a State-Contingent Conditional Order-m Approach

Teresa Serra[1], Alfons Oude Lansink[2]

[1] CREDA-UPC-IRTA; Parc Mediterrani de la Tecnologia; Esteve Terrades, 8; 08860 Castelldefels; Spain; e-mail: teresa.serra-devesa@upc.edu; Tel +34-93-552-1209; Fax: +34-93-552-1121

[2] Business Economics Group, Wageningen University; Hollandseweg, 1; 6706KN; Wageningen; The Netherlands; e-mail: alfons.oudelansink@wur.nl; Tel +31-317-485-194

Abstract

This article studies the influence of risk on farms' technical efficiency levels. The analysis extends the order-*m* efficiency scores approach proposed by Daraio and Simar (2005) to the state-contingent framework. The empirical application focuses on cross section data of Catalan specialised crop farms from the year 2011. Results suggest that accounting for production risks increases the technical performance. A 10% increase in output risk will result in a 2.5% increase in average firm technical performance.

Key words: linear programming, data envelopment analysis, nonparametric order-*m* efficiency, state-contingent technology, risk.

Paper prepared for presentation at the EAAE 2014 Congress
'Agri-Food and Rural Innovations for Healthier Societies'

August 26 to 29, 2014
Ljubljana, Slovenia

1. Introduction

Proper measurement of a firm's efficiency and productivity requires an appropriate characterization of the stochastic environment in which production takes place. Inadequate characterization of the stochastic environment may lead to uncertainty being incorrectly attributed to efficiency and productivity differences (Chambers and Quiggin, 2000; O'Donnell et al., 2010). Production under uncertainty can be characterized by differentiating outputs depending on the state of nature in which they are realized (Chambers and Quiggin, 1998 and 2000). This characterisation leads to a stochastic technology based on a state-contingent input correspondence. Conventional empirical representations of stochastic technologies are restrictive by requiring non-substitutability between state-contingent outputs. This restriction implies that producers can only respond to changes in the production environment by changing input use, but not reallocating production among different states of nature. Standard stochastic production functions are based upon the assumption that the state-contingent vector of outputs is generated by a vector of inputs that are controlled by the producer and a random variable beyond the control of the producer (Chambers and Quiggin, 2002, 2006). The state-contingent input correspondence approach developed by Chambers and Quiggin (2000) does not impose this restriction, as it assumes that the vector of state-contingent outputs is chosen ex-ante, though realized ex-post. The state-contingent framework has been shown to yield more precise efficiency measures relative to approaches that impose this restriction.

Previous literature has provided ample evidence that economic agents are not neutral to risk. To the extent that economic decisions are influenced by risk preferences, risk implicit in the state-contingent output distribution may have an impact on the efficiency with which economic agents operate (Battese et al., 1997). We measure this impact through Daraio and Simar's (2005) nonparametric frontier model, that allows for the influence of external factors on firm efficiency ratings. In frontier analysis, nonparametric efficiency measures are based upon the assumption that all observed units belong to the attainable production set. As a result, super-efficient outliers can have an influential impact on these envelopment estimators. Robustness can be increased through a trimming process that results in the frontier not enveloping all data points. Daraio and Simar (2005) provide a probabilistic formulation of a robust nonparametric order-m efficiency model, being m the trimming parameter, that allows for the influence of environmental variables that cannot be controlled by the producer, but that shape the outcome of production. Daraio and Simar (2005) proposal allows determining whether the environmental variable promotes or reduces efficiency. However, the Daraio and Simar (2005) approach does not adequately capture the stochastic conditions under which production takes place. The state-contingent framework proposed by Chambers and Quiggin (2000) can be implemented using standard tools of efficiency analysis when ex-ante outputs are known. As a result, Daraio and Simar's (2005) framework can be extended to examine efficiency and productivity in truly state-contingent terms. The extended model will not only be robust to outliers, but also to incorrect interpretations of uncertainty effects as efficiency effects.

The objective of this research article is to assess the influence of production risk on the technical efficiency of a sample of farms by extending the approach developed by Daraio and Simar to the state-contingent production technology. The empirical application focuses on cross-section data of arable farms in Catalonia, Spain. Eliciting information on ex-ante state-contingent outputs is a highly complex process that can be subject to subjectivity regarding beliefs on the crop yield distribution. Therefore, this paper generated order-*m* efficiency scores that are more robust to the presence of outliers. The efficiency scores were derived for each sample farm and the role of risk on farm performance was assessed.

2. Methods

Within the state-contingent framework, uncertainty is represented by a set of states of nature Ω from which nature makes a draw. Random variables in the production process can be measured as maps from the set of states Ω to the reals. Assume a single random output firm. The random output can be represented as a vector $\tilde{y} \in \mathbb{R}_+^{\Omega}$, where $\tilde{y} = \{y_s : s \in \Omega\}$, being y_s the realized (ex post) value of the random output variable $\tilde{y}$ if nature chooses state s. The non-random input vector is denoted by $x \in \mathbb{R}_+^N$. Denote by $Z \in \mathbb{Z} \subset \mathbb{R}^r$ the vector of environmental factors that are exogenous to the production process, but may explain part of it.

The stochastic production technology is represented by $\psi := \{(x, \tilde{y}) : x \text{ can produce } \tilde{y}\}$. The boundaries of ψ are an indicator of the efficiency with which firms operate. Under the influence of environmental variables, the technology is defined as $\psi^z := \{(x, \tilde{y}) | z : x \text{ can produce } \tilde{y}\}$. Note that for all $z \in \mathbb{Z}$, $\psi^z \subseteq \psi$. The interpretation of the technology is as follows: before knowing the realization of the state of nature, the producer chooses $(x, \tilde{y})$ from within the technology set, thus making a decision about nonstochastic inputs and stochastic outputs. After this selection has been made, nature makes a choice from $s \in \Omega$. For agricultural technologies, $s \in \Omega$ is usually related to weather conditions. It is important to note that ex-post realizations of random outputs are chosen by nature, and not by the producer (Chambers et al., 2011). Our article hypothesizes that the risk that firms face in the process from selecting the ex-ante output to obtaining ex-post realized production can have an impact on technical efficiency ratings, and we capture this risk through the environmental variable. Thus, in our particular application, $z = z(s)$.

Efficiency scores are usually approximated through radial distance from each production unit to the production frontier. Along these lines, the Farrell-Debreu output-oriented efficiency score for a firm operating with $\psi := \{(x, \tilde{y}) : x \text{ can produce } \tilde{y}\}$ can be defined as $\lambda(x, \tilde{y}) = \sup\{\lambda | (x, \lambda\tilde{y}) \in \psi\}$, being $\lambda(x, \tilde{y})$ the proportionate increase in outputs that can be achieved using the same technology and input combination. Since it is unknown, an estimator of the production frontier is required. Commonly used nonparametric estimators such as the Data Envelopment Analysis (DEA) initiated by Farrell (1957), or the Free Disposal Hull (FDH) proposed by Deprins et al. (1984) are based on the envelopment approach, which assumes that all observed units belong to the attainable set. While DEA and FDH have their own strengths and weaknesses, some scholarly papers have argued that FDH provides a better data fit than DEA (Tulkens, 1993; Vanden Eeckout et al., 1993). FDH usually outperforms DEA in technical efficiency measurement, because it constructs a technology that envelops the data more closely than DEA. Our analysis stems from the FDH technology set estimator which can be expressed as $\hat{\psi}_{FDH} := \{(x, \tilde{y}) \in \mathbb{R}_+^{N+\Omega} | \tilde{y} \le \tilde{Y}_i, x \ge X_i \; i = 1, \ldots, n\}$, where $i = 1, \ldots, n$

denotes the observation number. The empirical problem consists of estimating the frontier and the efficiency scores from a random sample of production units $\chi = \left\{ (X_i, Y_i) \middle| i = 1,...,n \right\}$.

Cazals et al. (2002) have proved that, under free disposability of inputs and outputs, a probability function of $(X, \tilde{Y})$ on $\mathbb{R}_+^N \times \mathbb{R}_+^\Omega$, $H(x, \tilde{y}) = \Pr(X \leq x, \tilde{Y} \geq \tilde{y})$, can be used to characterize the production frontier. $H(x, \tilde{y})$ represents the probability of dominating a unit that operates at the level $(x, \tilde{y})$ and can be decomposed as $H(x, \tilde{y}) = \Pr(\tilde{Y} \geq \tilde{y} \middle| X \leq x) \Pr(X \leq x) = S_{\tilde{Y}|X}(\tilde{y}|x) F_X(x)$. The survival function $S_{\tilde{Y}|X}(\tilde{y}|x)$ represents the attainable output set for a producer using no more than the input level x, and can be used to characterize the Farrell-Debreu output efficiency measure as follows:

$$\lambda(x, \tilde{y}) = \sup\left\{ \lambda \middle| S_{\tilde{Y}|X}(\lambda \tilde{y} | x) > 0 \right\}. \tag{1}$$

A nonparametric estimator of efficiency can be provided by plugging an empirical version of $S_{\tilde{Y}|X}(\tilde{y}|x)$ in (1):

$$\hat{\lambda}_n(x, \tilde{y}) = \sup\left\{ \lambda \middle| \hat{S}_{\tilde{Y}|X,n}(\lambda \tilde{y} | x) > 0 \right\}, \tag{2}$$

where n is the sample size and $\hat{S}_{\tilde{Y}|X,n}(\tilde{y}|x) = \dfrac{\sum_{i=1}^{n} \Gamma(X_i \leq x, \tilde{Y}_i \geq \tilde{y})}{\sum_{i=1}^{n} \Gamma(X_i \leq x)}$ is an estimator of $S_{\tilde{Y}|X}(\tilde{y}|x)$, being $\Gamma(.)$ an indicator function. Expression (2) has been shown by Cazals et al. (2012) to coincide with the FDH estimator of efficiency that is given by:

$$\hat{\lambda}_n(x, \tilde{y}) = \max_{i|X_i \leq x} \left\{ \min_{s=1,...,\Omega} \left(\frac{Y_{s,i}}{y_s} \right) \right\}, \tag{3}$$

where s denotes the state of nature.

FDH efficiency estimates are very sensitive to the presence of outliers (i.e., atypical observations substantially different from the rest of the data). By definition, nonparametric estimates of the production frontier are extreme values of the dimensional space of inputs and outputs. This implies that the presence of super-efficient outliers may significantly affect the shape of the FDH frontier and efficiency computations. Cazals et al. (2002) and Daraio and Simar (2005) propose a generalized version of the FDH nonparametric approach, the order- m frontier, that is more robust to extreme observations. The method uses, as a benchmark to evaluate a firm's performance, the expected value of the best practice among m peers randomly drawn from the population of firms that use input levels up to x. For a given input level x in the interior of the support of X, consider m i.i.d. random variables $\tilde{Y}_i$, $i = 1,...,m$ generated by the conditional distribution function $S_{\tilde{Y}|X}(\tilde{y}|x)$ and define $\Psi_m(x) = \left\{ (x', \tilde{y}) \in \mathbb{R}_+^{N+\Omega} \middle| x' \leq x, \tilde{Y}_i \geq \tilde{y}, i = 1,...,m \right\}$, where $\tilde{y}$ is generated though $S_{\tilde{Y}|X}(\tilde{y}|x)$. The maximum radial expansion for $(x, \tilde{y})$ to reach the FDH of the random set of firms $(x, \tilde{Y}_i)$ can be represented as:

$$\tilde{\lambda}_m(x, \tilde{y}) = \sup\left\{\lambda > 0 \middle| (\lambda \tilde{y} \middle| x) \in \Psi_m\right\} = \max_{i=1,\dots,m}\left\{\min_{s=1,\dots,\Omega}\left(\frac{Y_{s,i}}{y_s}\right)\right\}. \tag{4}$$

The estimator of the expected order-*m* efficiency score is then defined as:

$$\hat{\lambda}_m(x, \tilde{y}) = \mathrm{E}(\tilde{\lambda}_m(x, \tilde{y}) \big| X \le x) = \int_0^\infty \left[1 - (1 - \hat{S}_{\tilde{Y}|X,n}(u\tilde{y} \big| X \le x))^m\right] du. \tag{5}$$

Since the reference against which efficiency of a production unit is measured is the average of the best among *m* peers, the frontier is less extreme than the FDH. Noteworthy is the fact that the order-*m* efficiency score is not bounded by 1. Efficiency values greater than 1 indicate that the firm is more efficient than the average of randomly drawn *m* peers.

The model presented above can be extended to allow for exogenous factors that are not under the control of the firm, but that influence its performance levels. Under this setting, any random event is conditioned to $Z = z$. The joint distribution function of $(X, \tilde{Y})$ conditional on $Z = z$, is defined as: $H(x, \tilde{y} | z) = \Pr(X \le x, \tilde{Y} \ge \tilde{y} | Z = z)$ and its decomposition as $H(x, \tilde{y} | z) = \Pr(\tilde{Y} \ge \tilde{y} | X \le x, Z = z)\Pr(X \le x, Z = z) = S_{\tilde{Y}|X,Z}(\tilde{y} | x, z) F_{X|Z}(x | z)$. The Farrell-Debreu output efficiency estimate becomes:

$$\hat{\lambda}_n^z(x, \tilde{y} | z) = \sup\left\{\lambda \middle| \hat{S}_{\tilde{Y}|X,Z,n}(\lambda \tilde{y} | x, z) > 0\right\}, \tag{6}$$

where $\hat{S}_{\tilde{Y}|X,Z,n} = \dfrac{\sum_{i=1}^n \Gamma(X_i \le x, \tilde{Y}_i \ge \tilde{y}) K((Z_i - z)/h_n)}{\sum_{i=1}^n \Gamma(X_i \le x) K((Z_i - z)/h_n)}$ is an estimator of $S_{\tilde{Y}|X,Z}$, $\chi = \left\{(X_i, Y_i, Z_i) \middle| i = 1, \dots, n\right\}$ is a sample of *n* iid observations, $K((Z_i - z)/h_n)$ is a kernel function and h_n is the bandwidth. As shown by Daraio and Simar (2005), equation (6) coincides with the conditional FDH estimator of efficiency, that is given by:

$$\hat{\lambda}_n^z(x, \tilde{y} | z) = \max_{i | X_i \le x, |Z_i - z| \le h}\left\{\min_{s=1,\dots,\Omega}\left(\frac{Y_{s,i}}{y_s}\right)\right\}, \tag{7}$$

where *h* is the chosen bandwidth.

The conditional order-*m* frontier can be defined as $\Psi_m^z(x) = \left\{(x', \tilde{y}) \in \mathbb{R}_+^{N+\Omega} \middle| x' \le x, \tilde{Y}_i \ge \tilde{y}, i = 1, \dots, m\right\}$, where $\tilde{y}$ is generated through $S_{\tilde{Y}|X,Z}$. The maximum expansion for

$(x,\tilde{y})$ to reach the conditional FDH of the random set of firms $(x,\tilde{Y}_i)$ $i=1,...,m$, can be represented as:

$$\widehat{\lambda}_m^z(x,\tilde{y}) = \sup\left\{\lambda > 0 \middle| (\lambda\tilde{y}|x) \in \Psi_m^z\right\} = \max_{i=1,...,m}\left\{\min_{s=1,...,\Omega}\left(\frac{Y_{s,i}}{y_s}\right)\right\}. \tag{8}$$

The expected order-*m* efficiency score is defined as:

$$\hat{\lambda}_m(x,\tilde{y}|z) = \mathrm{E}(\hat{\lambda}_m(x,\tilde{y}|z) \big| X \le x) = \int_0^\infty \left[1-(1-\hat{S}_{\tilde{Y}|X,Z}(u\tilde{y}|X \le x, Z=z))^m\right] du. \tag{9}$$

Due to the multivariate nature of $S_{\tilde{Y}|X}$ and $S_{\tilde{Y}|X,Z}$, there are no simple expressions of $\hat{\lambda}_m(x,\tilde{y})$ and $\hat{\lambda}_m(x,\tilde{y}|z)$. As shown by Daraio and Simar (2005), the integrals in (5) and (9) can be evaluated using a simple Monte Carlo algorithm that is described below. First, for a given x, draw a sample of size *m* with replacement among those $\tilde{Y}_i$ such that $X_i \le x$ and denote the sample by $(\tilde{Y}_{1b},...,\tilde{Y}_{mb})$. Second, compute $\widehat{\lambda}_m^b(x,\tilde{y}) = \max_{i=1,...,m}\left\{\min_{s=1,...,\Omega}\left(\frac{Y_{s,i,b}}{y_s}\right)\right\}$. Third, repeat the first and second steps for $b=1,...,B$, where B is large. Finally, compute the average of the bootstrapped efficiency estimates: $\hat{\lambda}_{m,n}(x,\tilde{y}) \approx \frac{1}{B}\sum_{b=1}^{B}\widehat{\lambda}_m^b(x,\tilde{y})$. This Monte-Carlo algorithm changes slightly when we need to compute the conditional order-*m* efficiency score. Specifically, in the first stage, we need to draw a sample of size *m* with replacement and with a probability $\frac{K((z_i-z)/h)}{\sum_{j=1}^{n}K((z-z_j)/h)}$, among those $\tilde{Y}_i$ such that $X_i \le x$. The rest of the process does not change.

Conditional order-*m* efficiency estimators are based upon the smoothing in the estimation of function $\hat{S}_{\tilde{Y}|X,Z,n}(\lambda\tilde{y}|x,z)$. Our bandwidth selection process follows Daraio and Simar (2005) two stage approach. In the first stage, the likelihood cross-validation criterion, using a k-Nearest Neighbour method is used to optimize the estimation of the density function of Z. In a second step, the dimensionality of $\tilde{y}$ and x is considered in order to compute $\hat{S}_{\tilde{Y}|X,Z,n}(\lambda\tilde{y}|x,z)$, which is done by expanding the bandwidth obtained in the first stage by a factor of $1+n^{(-1/(N+\Omega))}$.

The impacts of Z on the production process are studied using a nonparametric regression that investigates the relationship between the ratio of the conditional and unconditional efficiency scores $\frac{\hat{\lambda}_{m,n}(x,\tilde{y}|z)}{\hat{\lambda}_{m,n}(x,\tilde{y})}$ and Z. More specifically, a local linear regression method is adopted to model dependence between these two variables. Details on the nonparametric linear regression methodology can be found in Fan and Gijbels (1996). Such nonparametric approach allows for non-monotonic impacts of exogenous variables on efficiency levels, i.e., it allows the impacts of these variables to change depending on the level of Z.

3. Data

A survey of 190 farms in Catalonia specialized in the production of arable crops was conducted in 2011, prior to the growing season. Farms were considered as specialized in arable crops if their arable crop income (cereals, oilseeds and protein crops) represented at least 80% of their global income. Data obtained from the survey included farm's planned input use (land, crop-specific inputs, farming overheads, paid and unpaid labor, and capital). Eight input variables were defined and used in the analysis: pesticide use expressed in total liters of active ingredients (x_1); organic and chemical fertilizers, expressed in kilos of nitrogen (x_2), total land planted to arable crops (measured in hectares, x_3), expenses in seeds (x_4), energy (x_5) and contract work in euros (x_6), total quantity of paid and unpaid labor (in hours, x_7), as well as the capital used in the production process (expressed as the replacement value in euros, x_8).

Data on the ex-ante cereal, oilseed and protein crop (COP) production were also collected. The elicitation process was designed as a trade-off between complexity and accuracy of responses. While providing survey respondents with detailed information on different crop-growing scenarios may increase response accuracy, the amount of information that one should provide would make the survey too long and time-consuming, as well as too expensive. For example, by indicating the number of frost days during the growing season, one would not be projecting a detailed enough scenario, as frost days are not equally damaging during the plant growth cycle. Technicians from Unió de Pagesos, the largest farmer association in Catalonia and the group in charge of conducting the survey, recommended obtaining point estimates under bad, normal and ideal growing conditions without projecting specific scenarios.

As argued by Unió de Pagesos, yields under normal growing conditions are usually a reference to producers, who typically provide, as a response, an average yield over a sufficiently long period of time (10 years approximately). Once these yields are identified, it is relatively easy for farmers to provide yield data under bad and ideal conditions. Three output variables were thus created: output under bad, normal and ideal growing conditions (y_1, y_2, y_3) and they were measured in euros, using the expected price under normal market conditions. As noted above, while the producer picks the vector (x, y_1, y_2, y_3) from within the technology set, nature has the role to choose the output that is being realized ex-post. Output risk is approximated by the standard deviation of the ex-ante output (i.e., output under the different crop growing conditions) expressed on a per hectare basis. It is important to note that our data collection method provides an estimate of the farmers' perception of the production function, and not of the actual state-contingent production function as this is not observed. A subjective estimate of the magnitude of output standard deviation is known by producers at planting time. Table 1 presents summary statistics for the variables of interest.

The difficulties associated to measuring ex-ante outputs are relevant. Our survey obtains these based on subjective notions that create potential for identification biases. For example, one farmer's normal state may be another's bad state. Difficulties inherent in obtaining accurate estimates of state-contingent yields are comparable to the hypothetical-bias problem in contingent-valuation analyses. This leads to differences in values drawn up under a hypothetical and a real setting (List and Shogren, 1998). In order to reduce this bias, calibration techniques have been proposed. Some of these techniques involve warning survey respondents about the quality of the responses. This is the case of the "cheap-talk" technique that has been shown to reduce unknowledgeable respondents' bias (List, 2001; Cummings and Taylor, 1999), but to be less successful with experienced respondents (List, 2001; Lusk, 2003). Our respondents had 29 years of experience in cultivating arable crops, and thus they can be considered as experienced. The fact that interviewers were technicians from Unió de Pagesos who know the farm well is expected to have reduced farmers' incentives to provide biased responses.

Table 1. Descriptive statistics of the variables used in the empirical application

Variable	Variable Description	Mean	Standard deviation	Minimum	Maximum
y_1	Output under bad growing conditions (€)	30,444.95	33,117.37	1,630.20	246,000.00
y_2	Output under normal growing conditions (€)	50,737.46	51,625.90	3,003.00	396,000.00
y_3	Output under ideal growing conditions (€)	70,111.90	74,725.74	3,912.48	576,000.00
x_1	Active ingredients applied (liters)	84.89	117.05	0.00	749.70
x_2	Nitrogen application (kg)	9,658.31	12,217.45	82.44	95,085.47
x_3	Land (ha)	74.54	72.59	7.51	500.00
x_4	Seeds (€)	3,851.66	3,745.45	104.89	24,800.00
x_5	Energy (€)	4,891.20	5,328.66	0.00	41,250.00
x_6	Contract work (€)	2,901.19	4,013.90	0.00	26,820.00
x_7	Labour (hours)	549.90	655.42	0.00	3,775.96
x_8	Capital replacement value (€)	133,453.09	126,276.96	0.00	813,000.00
z	Production risk	263.75	93.29	70.96	603.99

4. Results

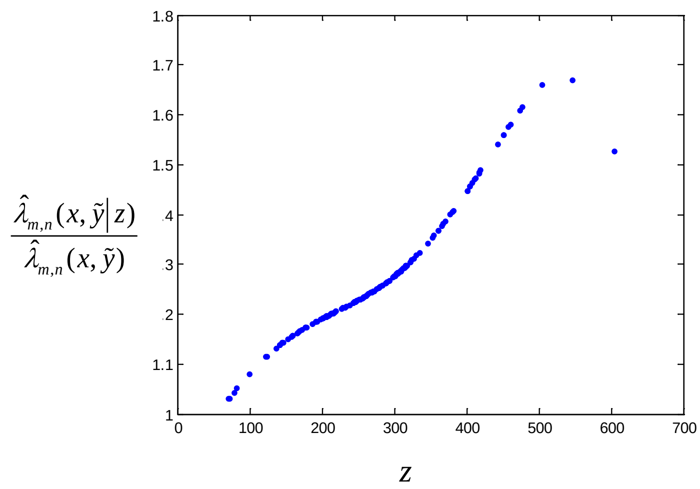
In the application of the order-m efficiency model, a trimming parameter of the frontier equal to *m*=10 was selected. The results of the estimation of unconditional $\left(\hat{\lambda}_{m,n}(x,\tilde{y})\right)$ and production risk conditional $\left(\hat{\lambda}_{m,n}(x,\tilde{y}\,|\,z)\right)$ order-*m* technical efficiency are found in Table 2. The average unconditional and conditional technical efficiency scores take values of 0.84 and 0.98, respectively. This suggests, first, that producers in the sample generate 16% and 2% less output than technically feasible in the unconditional and conditional setting. Second, these results show that consideration of production risk yields higher technical efficiency scores, i.e. that risk has a positive effect on technical performance. Hence, the more risky the production environment, the higher the motivation to produce efficiently. By improving technical performance in response to a more risky production environment, farmers build up larger buffers in good states of nature and are better prepared for the bad states of nature. Hence, by improving the performance, farmers compensate for the higher risks. These results are compatible with Chambers et al. (2011) and Skevas et al. (2012) who find that higher risk levels lead to better technical performance.

Table 2. Frequency distribution of Daraio and Simar (2005) order-m efficiency scores

Efficiency interval	order-*m* efficiency estimates	
	$\hat{\lambda}_{m,n}(x,\tilde{y})$	$\hat{\lambda}_{m,n}(x,\tilde{y}\mid z)$
$0 < \lambda < 0.1$	0	0
$0.1 \le \lambda < 0.2$	0	0
$0.2 \le \lambda < 0.3$	1	0
$0.3 \le \lambda < 0.4$	5	0
$0.4 \le \lambda < 0.5$	11	0
$0.5 \le \lambda < 0.6$	15	0
$0.6 \le \lambda < 0.7$	16	2
$0.7 \le \lambda < 0.8$	19	5
$0.8 \le \lambda < 0.9$	21	8
$0.9 \le \lambda$	102	175
Mean score	**0.8380**	**0.9786**

The positive influence of risk on technical efficiency is formalized through the smoothed nonparametric regression line of $\dfrac{\hat{\lambda}_{m,n}(x,\tilde{y}\mid z)}{\hat{\lambda}_{m,n}(x,\tilde{y})}$, the efficiency ratio, over z presented in Figure 1. The marginal impact of risk on the efficiency ratio is not constant: Figure 2 shows that for risk levels up to 200 euros per hectare, the marginal impact of production risk on technical efficiency is positive but declining. Hence, below risk levels of 200 euros per hectare an additional euro of risk increases technical efficiency, although the marginal improvement becomes smaller. However, for risk levels above 200 euro per hectare, the marginal impact on performance sharply increases with risk before levelling off at the risk level of almost 500 euro per hectare. This result suggests that, the higher the production risk perceived by farmers, the higher their effort to improve their technical efficiency. However, at very high risk levels, i.e. risk levels beyond 500 euros per hectare, farmers do not put effort into further improving their technical efficiency. This result may imply that farmers have reached the limit of technical efficiency improvement, or concluded that further efforts into improving efficiency are not economically feasible.

Figure1. Smoothed local linear regression: effect of z on $\dfrac{\hat{\lambda}_{m,n}(x,\tilde{y}|z)}{\hat{\lambda}_{m,n}(x,\tilde{y})}$



In order to get a better grasp of the relevance of the magnitudes in Figures 1 and 2, the elasticity of ratio $\dfrac{\hat{\lambda}_{m,n}(x,\tilde{y}|z)}{\hat{\lambda}_{m,n}(x,\tilde{y})}$ with respect to Z has been computed at each data point. Results are presented in Figure 3. The average elasticity takes the value of 0.25, which implies that a 10% increase in output risk will result in a 2.5% increase in firm technical performance. As can be appreciated, there is an increase in elasticity as output risk increases, which suggests that the bigger the production risk, the greater the impact on efficiency changes. While average elasticity is 0.25, risk levels on the order of 500 euros involve elasticities around to 0.9.

Conditional efficiency estimates depend on the choice of the bandwidth through the likelihood cross validation criteria as described in the methodological section. Table 3 presents a sensitivity analysis to the choice of the bandwidth along the lines of Daraio and Simar (2005). Conditional order-*m* efficiency scores are computed for $h'=h*1.2$ and $h'=h*0.8$, where h is the bandwidth determined by the data-driven method. Mean values are presented in the table. Results are remarkably stable to the bandwidth choice. While not presented in the paper, but available upon request, the shapes of Figures 1 and 2 do not change with the bandwidth selection. Table 3 further shows sensitivity of efficiency scores to changes to the value of *m* (i.e., the number of random units or peers considered in the definition of the technology set). Results are robust to the change in the size of the peer group.

Figure 2. Smoothed local linear regression: marginal impact of z on $\dfrac{\hat{\lambda}_{m,n}(x,\tilde{y}|z)}{\hat{\lambda}_{m,n}(x,\tilde{y})}$

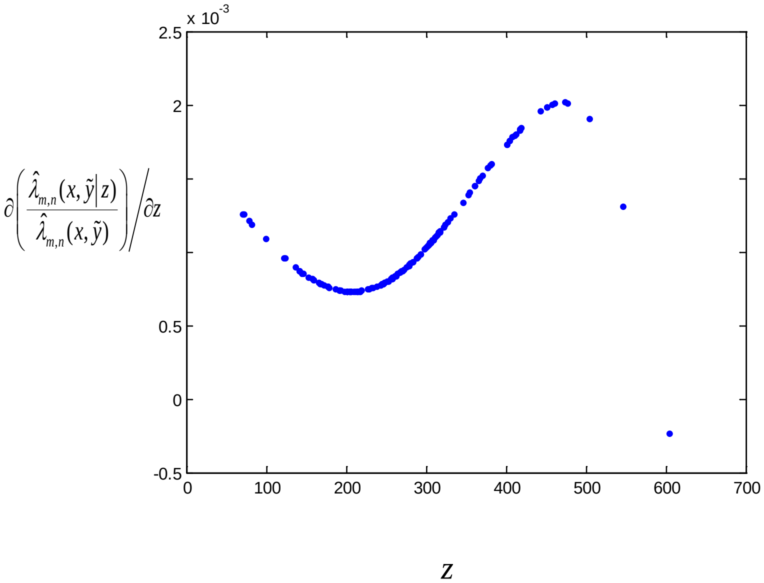


Figure 3. Elasticity of $\dfrac{\hat{\lambda}_{m,n}(x,\tilde{y}|z)}{\hat{\lambda}_{m,n}(x,\tilde{y})}$ with respect to the environmental variable z

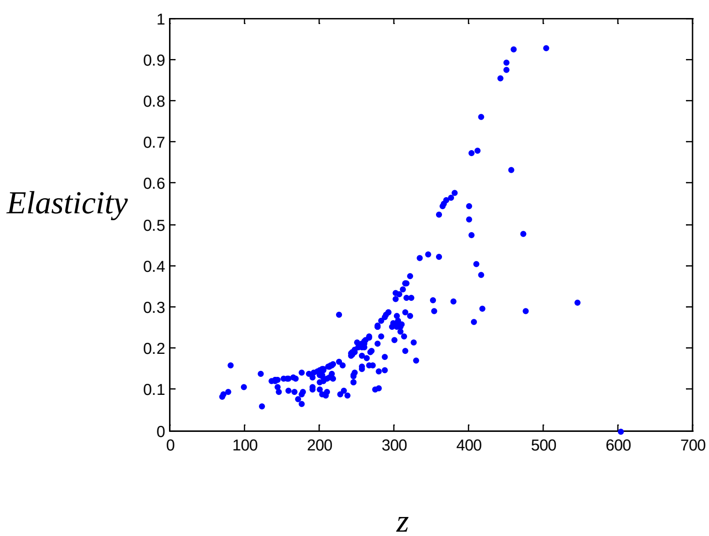


Table 3. Efficiency estimates sensitivity to the choice of the bandwidth and to the choice of the order of the frontier

	$\hat{\lambda}_{m,n}(x,\tilde{y}\|z)$
$h'=h*0.8$	0.9837
$h'=h$	0.9786
$h'=h*1.2$	0.9731
$m=10$	0.9786
$m=20$	0.9944
$m=30$	0.9983

Note: *h* is the bandwidth determined by the data-driven method

4. Concluding remarks

This article studies the influence of risk on farmers' technical efficiency. We use nonparametric order-*m* efficiency scores allowing for the influence of environmental variables as proposed by Daraio and Simar (2005). While we recognize that other variables may affect efficiency ratings, the curse of dimensionality of nonparametric estimators recommends against increasing the model size. The Daraio and Simar (2005) framework is extended to examine efficiency in a state-contingent production technology. The use of the state-contingent framework to account for the stochastic conditions under which production takes place entails an improvement compared to conventional approaches. Previous research has shown that the use of the state-contingent approach leads to enhanced efficiency estimates. Results suggest that the risk perception by economic agents increases the efficiency with which they operate. More specifically, an increase in output risk on the order of 10% will result in a 2.5% increase in firm technical performance. Nonparametric regression analysis of the influence of risk on efficiency ratings indicates that, while relatively low efficiency levels motivate slow increases in efficiency, high risk levels accelerate improvements in farm performance. A direction for future extensions of this work lies in comparing our results with robust efficiency measures derived using bootstrapping techniques along the lines of Simar and Wilson (2000).

The authors gratefully acknowledge financial support from Instituto Nacional de Investigaciones Agrícolas (INIA) and the European Regional Development Fund (ERDF), Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+i), Project Reference Number RTA2012-00002-00-00.

References

Battese, G.E., Rambaldi, A.N., Wan, G.H. (1997). A stochastic frontier production function with flexible risk properties. *Journal of Productivity Analysis*, 8, 269–280.

Cazals, C., Florens, J.P., Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106, 1-25.

Chambers, R.G., Hailu, A., Quiggin, J. (2011). Event-specific data envelopment models and efficiency analysis. *Australian Journal of Agricultural and Resource Economics*, 55, 90-106.

Chambers, R.G., Quiggin, J. (1998). Cost functions and duality for stochastic technologies. *American Journal of Agricultural Economics*, 80, 288-295.

Chambers, R.G., Quiggin, J. (2000). *Uncertainty, production, choice, and agency*. Cambridge: Cambridge University Press.

Chambers, R.G., Quiggin, J. (2002). The state-contingent properties of stochastic production functions. *American Journal of Agricultural Economics*, 84, 513-526.

Chambers, R.G., Quiggin, J. (2006). The state-contingent approach to production under uncertainty. *Australian Journal of Agricultural and Resource Economics*, 50, 153-159.

Cummings, R.G., Taylor, L.O. (1999). Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review*, 89,649-665.

Daraio, C., Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: a probabilistic approach. *Journal of Productivity Analysis*, 24, 93-121.

Deprins, D., Simar, L., Tulkens, H. (1984). Measuring labor-efficiency in post offices. In M. Marchand, P. Pestieau and H. Tulkens (Eds.), *The performance of public enterprises - Concepts and measurement* (pp. 243-267). Amsterdam: North Holland.

Fan, J., Gijbels, I. (1996). *Local polynomial modeling and its applications*. London: Chapman & Hall.

Farrell, M.J. (1957). The measurement of the productive efficiency. *Journal of the Royal Statistical Society*, Series A, CXX, Part 3, 253-290.

List, J.A. (2001). Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sport cards. *American Economic Review*, 91, 1498-1507.

List, J.A., Shogren, J.F. (1998). Calibration of the differences between actual and hypothetical valuations in a field experiment. *Journal of Economic Behavior and Organization*, 37,193-205.

Lusk, J.L. (2003). Effects of cheap talk on consumer willingness-to-pay for golden rice. *American Journal of Agricultural Economics*, 85, 840-856.

Nelson, C.H., Preckel, P.V. (1991). The conditional beta distribution as a stochastic production function. *American Journal of Agricultural Economics*, 2,370-378.

O'Donnell, C.J., Chambers, R.G., Quiggin, J. (2010). Efficiency analysis in the presence of uncertainty. *Journal of Productivity Analysis*, 33, 1-17.

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

Skevas, T., Oude Lansink, A., Stefanou, S. (2012). Measuring technical efficiency in the presence of pesticide spillovers and production uncertainty: the case of Dutch arable farms. *European Journal of Operational Research*, 2, 550-559.

Tulkens, H. (1993). On FDH efficiency analysis: some methodological issues and applications to retail banking, courts and urban transit. *Journal of Productivity Analysis*, 4, 183-210.

Vanden Eeckout, P., Tulkens, H., Jamar, M.A. (1993). Cost efficiency in Belgian municipalities. In H.O. Fried, C.A.K. Lovell, and S.S. Schmidt (Eds.), *The Measurement of productive efficiency: Techniques and applications* (pp. 330-334). New York: Oxford University Press.