



## Production under a quota system: an extension of the weighted additive model to assess technical efficiency

Juan Aparicio, Magdalena Kapelko, Lidia Ortiz & Jesus T. Pastor

To cite this article: Juan Aparicio, Magdalena Kapelko, Lidia Ortiz & Jesus T. Pastor (2017): Production under a quota system: an extension of the weighted additive model to assess technical efficiency, INFOR: Information Systems and Operational Research, DOI: [10.1080/03155986.2016.1272961](https://doi.org/10.1080/03155986.2016.1272961)

To link to this article: <http://dx.doi.org/10.1080/03155986.2016.1272961>



Published online: 03 Jan 2017.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



# Production under a quota system: an extension of the weighted additive model to assess technical efficiency

Juan Aparicio<sup>a</sup>, Magdalena Kapelko<sup>b</sup>, Lidia Ortiz<sup>a</sup> and Jesus T. Pastor<sup>a</sup>

<sup>a</sup>Center of Operations Research (CIO), Miguel Hernandez University of Elche (UMH), Elche (Alicante), Spain;

<sup>b</sup>Department of Logistics, Institute of Applied Mathematics, Wrocław University of Economics, Wrocław, Poland

## ABSTRACT

Since it was introduced, data envelopment analysis (DEA) has been applied to many different areas and has also been extended to numerous production contexts. In this paper, we focus our attention on the production framework under output quotas. Many types of markets throughout the world are subject to intervention by government policies, with the purpose of raising the prices that producers receive for their products through the imposition of production quotas. In this paper, we introduce a comprehensive approach based upon the extension of the weighted additive model in DEA with the aim of measuring technical inefficiency of firms producing under a quota system. Accordingly, we prove several interesting properties of the new approach and compare it with the conventional weighted additive model. The paper concludes by considering a case study that analyses milk production inefficiency in Canadian provinces, incorporating information on the quota assigned to each of the provinces.

## ARTICLE HISTORY

Received 28 September 2016

Accepted 9 December 2016

## KEYWORDS

Data envelopment analysis; the weighted additive model; production under quotas

## 1. Introduction

Data envelopment analysis (DEA) is a non-parametric method for the assessment of relative technical efficiency of a set of decision making units (DMUs) that use several inputs to produce several outputs. Technical efficiency refers to all the sources of waste that can be removed without worsening any input and/or output. In addition to the efficiency estimation of each observed DMU, DEA also provides benchmarking information, which is useful for improving the performance of any inefficient DMU.

Chronologically speaking, Farrell (1957) was the first in showing, for a single output and multiple inputs, how to estimate an isoquant enveloping all the observations and how to measure technical efficiency. Farrell inspired other authors to continue this line of research estimating production functions that envelop all the observations of the sample, by either a non-parametric piece-wise linear technology or a parametric function. The first possibility was taken up by Charnes et al. (1978) and Banker et al. (1984), resulting in the development of DEA, whereas the latter approach resulted in the development of the deterministic and stochastic frontier models through parametric techniques.

In contrast to the parametric literature on efficiency, where the measurement of technical efficiency in the context of multiple-outputs is based on a few measures (the Shephard input and output distance functions and the directional distance function), the first years of life of DEA witnessed the introduction of many different technical efficiency measures, such as the Russell input and output measures of technical efficiency and their graph extension, that is the Russell Graph Measure of technical efficiency (see Färe et al. 1985), the additive model (Charnes et al. 1985), the Range-Adjusted Measure (Cooper et al. 1999) and the Enhanced Russell Graph (Pastor et al. 1999) or Slacks-Based Measure (Tone 2001), to name but a few. One of the reasons for the introduction of many different technical efficiency measures in DEA is the piece-wise linear nature of the boundary of the corresponding polyhedral technology. In this context, a notion that comes into play is Pareto-efficiency (Koopmans 1951). Pareto-efficiency, however, does not appear to be a problem for the parametric approach, where the functional forms utilized to model the frontier of production are generally smooth. On the contrary, it has been a recurring theme in DEA. In particular, the additive model by Charnes et al. (1985) was the first graph linear model which ensured that the evaluated DMU was compared exclusively with respect to the set of Pareto-efficient points in the input–output space. From this model, DEA researchers have introduced some modifications of the original additive model, weighting the slacks that appear in the objective function (see, for example, Lovell & Pastor 1995; Cooper et al. 1999, 2011), in order to measure technical inefficiency using the strongly efficient frontier as a reference. The existence of such a wide range of tools for estimating technical efficiency in the non-parametric world, in comparison with the parametric one, reveals the importance in DEA in measuring efficiency with respect to the Pareto-efficient frontier.

Since it was introduced, DEA has been applied to many different areas and has been also extended to numerous production contexts. In this paper, we particularly focus on the framework of production quotas. Many types of markets throughout the world are often subject to intervention by government policies with the purpose of raising the prices that producers receive for their products. Usually, this new price is above the equilibrium price that would have occurred in an unregulated market, which causes an excess supply and market forces will want to push the prices back down to the equilibrium level. There are generally two ways that government policies work against these market forces: one approach operates on the demand side and the other on the supply side. We will concentrate on the latter. Policies that operate on the supply side are called supply management policies because they set limits on production through quotas, thereby preventing an excess being produced. One example of this policy is the Canadian dairy industry quota system (Bunting 2009), which will be the focus of attention in the empirical section of this paper.

In view of these facts, the main aim of this paper is to introduce a new approach to modelling a quota production system, in order to evaluate technical inefficiency in this kind of regulated markets. The new approach will allow us to compare the level of technical inefficiency of each DMU under production limitations with a quota-free framework. The new model will be an extension of the weighted additive model in DEA, which introduces a comprehensive approach within a quota framework and ensures Pareto-efficiency.

In [Section 2](#), we are going to briefly review the ‘traditional’ weighted additive model in DEA. In [Section 3](#), we are going to propose a new approach for dealing with output quotas based on an extension of the weighted additive model, followed, in [Section 4](#), by an

empirical illustration for milk production under Canada's dairy quota system. Section 5 concludes the paper.

## 2. The weighted additive model in DEA

In this section, we briefly review the definition and main properties of the weighted additive model in DEA. Nevertheless, before doing so, we need to introduce the notation followed in the paper.

Let us consider  $n$  DMUs to be evaluated. DMU $_j$  consumes  $x_j = (x_{1j}, \dots, x_{mj}) \in R^m_+$  amount of inputs for the production of  $y_j = (y_{1j}, \dots, y_{sj}) \in R^s_+$  amount of outputs. The relative efficiency of each DMU in the sample is assessed with reference to the so-called production possibility set, which can be non-parametrically constructed from the observations by assuming certain postulates (see Banker et al. 1984). Also, to implement the approach, we will hereafter assume Variable Returns to Scale. In this way, the production possibility set in DEA,  $T$ , can be mathematically characterized as follows:

$$T = \{(x, y) \in R^{m+s}_+ : x \geq \sum_{j=1}^n \lambda_j x_j, y \leq \sum_{j=1}^n \lambda_j y_j, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n\}. \quad (1)$$

In the production literature, we can find the frontier concept linked to the notion of technology. Specifically, the weakly efficient frontier of  $T$  is defined as  $\partial^w(T) := \{(x, y) \in T : \hat{x} < x, \hat{y} > y \Rightarrow (\hat{x}, \hat{y}) \notin T\}$ , while the strongly efficient frontier of  $T$  is  $\partial^s(T) := \{(x, y) \in T : \hat{x} \leq x, \hat{y} \geq y, (\hat{x}, \hat{y}) \neq (x, y) \Rightarrow (\hat{x}, \hat{y}) \notin T\}$ , whereby the latter is a subset of the former. In words,  $\partial^s(T)$  is the set of all the non-dominated points of  $T$  in the sense of Pareto (see Koopmans 1951).

To estimate technical inefficiency for DMU $_0$  with data  $(x_0, y_0)$ , one possibility is to solve the following weighted additive model (Lovell & Pastor 1995):

$$\begin{aligned} \text{WA}(x_0, y_0, w^-, w^+) = & \quad \text{Max} \quad \sum_{i=1}^m w_i^- s_{i0}^- + \sum_{r=1}^s w_r^+ s_{r0}^+ \\ & \quad \text{s.t.} \\ & \quad \sum_{j=1}^n \lambda_{j0} x_{ij} = x_{i0} - s_{i0}^-, \quad i = 1, \dots, m \\ & \quad \sum_{j=1}^n \lambda_{j0} y_{rj} = y_{r0} + s_{r0}^+, \quad r = 1, \dots, s, \\ & \quad \sum_{j=1}^n \lambda_{j0} = 1, \\ & \quad s_{i0}^- \geq 0, \quad i = 1, \dots, m \\ & \quad s_{r0}^+ \geq 0, \quad r = 1, \dots, s \\ & \quad \lambda_{j0} \geq 0, \quad j = 1, \dots, n \end{aligned} \quad (2)$$

where  $w^- = (w_1^-, \dots, w_m^-) \in R^m_+$  and  $w^+ = (w_1^+, \dots, w_s^+) \in R^s_+$  are weights representing the relative importance of unit inputs and unit outputs, and  $s_{i0}^-$  and  $s_{r0}^+$  are the slacks associated with input  $i$  and output  $r$ , respectively.  $\text{WA}(x_0, y_0, w^-, w^+) = \sum_{i=1}^m w_i^- s_{i0}^{-*} + \sum_{r=1}^s w_r^+ s_{r0}^{+*}$ , where  $*$  denotes optimality, representing the technical

inefficiency associated with  $DMU_0$ . The weighted additive model maximizes a weighted  $\ell_1$  distance from  $DMU_0$  to the efficient frontier, and thereby simultaneously increases outputs and reduces inputs. Additionally, the targets of the evaluated unit, i.e. the coordinates of the projection point on the frontier, are defined by  $x_{i0}^* = \sum_{j=1}^n \lambda_{j0}^* x_{ij}$ ,  $i = 1, \dots, m$ , and  $y_{r0}^* = \sum_{j=1}^n \lambda_{j0}^* y_{rj}$ ,  $r = 1, \dots, s$ .

Next, we show two important properties of the weighted additive model that will be readapted later. Proofs for these results can be found, for example, in Pastor et al. (1999).

**Proposition 2.1:**  $(x_0, y_0) \in \partial^s(T)$  if and only if  $WA(x_0, y_0, w^-, w^+) = 0$ .

**Proposition 2.2:** Let  $(\lambda_0^*, s_0^{-*}, s_0^{+*})$  be an optimal solution of Equation (2). Then  $(x_0^*, y_0^*)$  belongs to  $\partial^s(T)$ .

The first result shows that the weighted additive model is able to characterize the Pareto-efficient frontier, whereas the second result establishes that the target point determined by programme (2) is always on the strongly efficient frontier. This last feature contrasts to the usual performance of other measures, such as those derived from the radial models or the directional distance function.

### 3. A weighted additive model for dealing with production quotas

In this section, we introduce an extension of the weighted additive model in order to incorporate production limitations linked to the existence of output quotas. To do that, we are going to consider a real production case where one output can be bounded by the requirement of the policy-maker. The case of multiple upper-bounded outputs can be similarly treated. In particular, we will focus our analysis on milk production with quota impositions in the rest of this section. First, we develop the basic model for production under quota and then we focus our attention on the improvement of this basic model that allows for additional properties to be satisfied.

#### 3.1. The formulation of a basic model for production with quotas

Let us assume that each DMU (dairy farm) produces two outputs, milk and cattle, from the consumption of several inputs. Let us assume, without loss of generality, that milk represents the first output. In addition, let us suppose that the market regulator imposes unit-specific milk quotas that must be taken into account. This is, for example, the case of Canada, where milk production quotas are allocated to each dairy farm and are bought and sold on an exchange (Bunting 2009). It is worth adding that the quota regime also existed in the dairy sector of European Union countries until very recently (2015 was the year of quota elimination).

In this context, the production possibility set for  $DMU_0$  with quota  $q_0$  is defined as

$$T^{q_0} = T \cap \{(x, y) \in R_+^{m+s} : y_1 \leq q_0\}, \quad (3)$$

which in the DEA setting would be translated into

$$T^{q_0} = \{(x, y) \in R_+^{m+s} : (x, -y) \geq \sum_{j=1}^n \lambda_j(x_j, -y_j), \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, \forall j, y_1 \leq q_0\}. \quad (4)$$

Therefore, by definition, the bounded technology  $T^{q_0}$  for  $(x_0, y_0)$  is the original ‘unbounded’ technology (1), adding a unit-specific upper bound for the milk produced.

In this quota-driven context, let us now show that if the researchers resort to a direct preadaptation of the ‘traditional’ weighted additive model (2) in order to determine technical inefficiency, by incorporating the new constraint  $y_1 \leq q_0$ , the resulting programme can present some weaknesses, such as identifying projections that are not Pareto-efficient. Under these hypotheses, the ‘natural’ modification of Equation (2) would be Equation (5), where a constraint for the output projection,  $y_0^* = \sum_{j=1}^n \lambda_j^* y_j$ , must be added in order to satisfy the quota constraint (5.4).

$$\begin{aligned} \text{Max} \quad & \sum_{i=1}^m w_{i0}^- s_{i0}^- + \sum_{r=1}^s w_{r0}^+ s_{r0}^+ \\ \text{s.t.} \quad & \sum_{j=1}^n \lambda_j x_{ij} = x_{i0} - s_{i0}^-, \quad i = 1, \dots, m \quad (5.1) \end{aligned}$$

$$\sum_{j=1}^n \lambda_j y_{rj} = y_{r0} + s_{r0}^+, \quad r = 1, \dots, s \quad (5.2)$$

$$\sum_{j=1}^n \lambda_j = 1, \quad (5.3)$$

$$\sum_{j=1}^n \lambda_j y_{1j} \leq q_0, \quad (5.4)$$

$$s_{i0}^- \geq 0, \quad i = 1, \dots, m \quad (5.5)$$

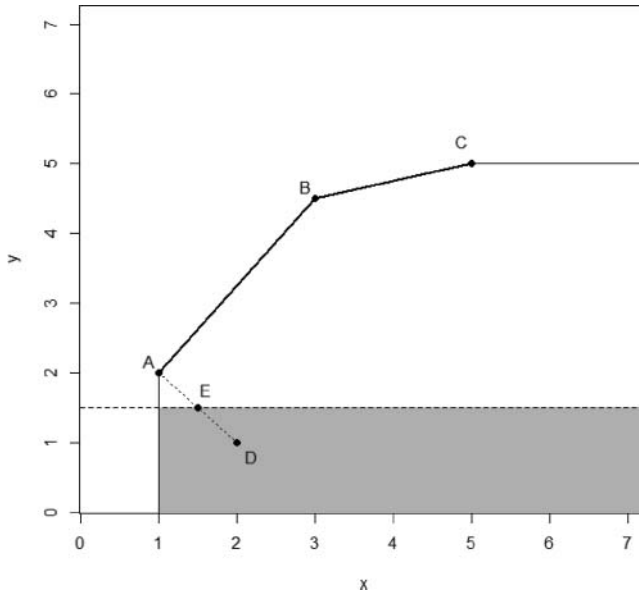
$$s_{r0}^+ \geq 0, \quad r = 1, \dots, s \quad (5.6)$$

$$\lambda_j \geq 0 \quad j = 1, \dots, n \quad (5.7)$$

Unfortunately, Equation (5) can yield inconsistent results regarding the essence of the weighted additive models, where the notion of Pareto-efficiency plays a key role. The weighted additive model has exactly the same behaviour as the additive model: the projections are always Pareto-efficient points (see Proposition 2.2). These ideas are graphically illustrated in Figure 1, based on a simple example with one input and one output. In this example, unit D is technically inefficient. Indeed, it is an interior point of the technology  $T$ . Let us assume that, for D, the output quota is  $q_D = 1.5$ . Note then that if unit D is assessed by means of model (5), with  $w_D^- = w_D^+ = 1$ , the optimal solution projects D onto E, a ‘dominated’ point of  $T^{q_D}$ , represented by the shaded area in the figure. In this respect, observe that the point  $(1, 1.5) \in T^{q_D}$  dominates E. Consequently, the a-priori ‘natural’ adaptation of the original weighted additive model to situations where output quotas are imposed, can lead to unsatisfactory results because Pareto-optimality is not achieved.

Fortunately, we may introduce a new approach to the production system described earlier and determine a suitable model for measuring technical inefficiency through a weighted additive measure that, in particular, will yield point  $(1, 1.5)$  as the target point for unit D in Figure 1. This point is Pareto-efficient for the technology associated to point D,  $T^{q_D}$ , although it is only weakly efficient in the original unbounded technology  $T$ .

The basic elements for building DEA models are, in general, the technology  $T$ , the corresponding distance function to the frontier for measuring technical inefficiency, the



**Figure 1.** Numerical example.

inputs and outputs of the observation being assessed  $(x_0, y_0)$ , and finally, the projection point  $(x_0^P, y_0^P)$ , i.e. the point that belongs to the frontier of the technology. In our context,  $T$  is substituted by  $T^{q_0}$ , and the distance function corresponds to the weighted  $\ell_1$  distance from  $(x_0, y_0)$  to  $(x_0^P, y_0^P)$ . The maximization of this distance will guarantee that the final identified projection belongs to the strongly efficient frontier of  $T^{q_0}$ , denoted as  $\partial^s(T^{q_0})$ , as we will prove later. In this way, the model would be written as

$$\begin{aligned}
 \text{Max} \quad & \sum_{i=1}^m w_{i0}^- s_{i0}^- + \sum_{r=1}^s w_{r0}^+ s_{r0}^+ \\
 \text{s.t.} \quad & x_{i0}^P = x_{i0} - s_{i0}^-, \quad i = 1, \dots, m \\
 & y_{r0}^P = y_{r0} + s_{r0}^+, \quad r = 1, \dots, s \\
 & (x_0^P, y_0^P) \in T^{q_0} \\
 & s_{i0}^- \geq 0, \quad i = 1, \dots, m \\
 & s_{r0}^+ \geq 0, \quad r = 1, \dots, s
 \end{aligned} \tag{6}$$

Going back to Equation (4), we can identify the particular structure of the projection point  $(x_0^P, y_0^P)$ , which belongs to  $T^{q_0}$ , in terms of the set of sample points as follows:

$$\begin{aligned}
 x_{i0}^P &= \sum_{j=1}^n \lambda_{j0} x_{ij} + \tau_{i0}^-, \quad \tau_{i0}^- \geq 0, \quad i = 1, \dots, m; \quad y_{r0}^P = \sum_{j=1}^n \lambda_{j0} y_{rj} - \tau_{r0}^+, \quad \tau_{r0}^+ \geq 0, \quad r = 1, \dots, s, \\
 \sum_{j=1}^n \lambda_{j0} y_{1j} - \tau_{10}^+ &\leq q_0, \quad \sum_{j=1}^n \lambda_{j0} = 1, \quad \lambda_{j0} \geq 0, \quad j = 1, \dots, n.
 \end{aligned} \tag{7}$$

As a consequence, Equation (6) may be rewritten as

$$\text{WA}^{q_0}(x_0, y_0, w^-, w^+) = \text{Max} \quad \sum_{i=1}^m w_{i0}^- s_{i0}^- + \sum_{r=1}^s w_{r0}^+ s_{r0}^+$$

$$\text{s.t.} \quad \sum_{j=1}^n \lambda_{j0} x_{ij} + \tau_{i0}^- = x_{i0} - s_{i0}^-, \quad i = 1, \dots, m \quad (8.1)$$

$$\sum_{j=1}^n \lambda_{j0} y_{rj} - \tau_{r0}^+ = y_{r0} + s_{r0}^+, \quad r = 1, \dots, s \quad (8.2)$$

$$\sum_{j=1}^n \lambda_{j0} y_{1j} - \tau_{10}^+ \leq q_0, \quad (8.3)$$

$$\sum_{j=1}^n \lambda_{j0} = 1, \quad (8.4)$$

$$\lambda_{j0} \geq 0, \quad j = 1, \dots, n \quad (8.5)$$

$$s_{i0}^- \geq 0, \quad i = 1, \dots, m \quad (8.6)$$

$$s_{r0}^+ \geq 0, \quad r = 1, \dots, s \quad (8.7)$$

$$\tau_{i0}^- \geq 0, \quad i = 1, \dots, m \quad (8.8)$$

$$\tau_{r0}^+ \geq 0, \quad r = 1, \dots, s \quad (8.9)$$

Note that the main difference between model (8) and model (5) is the incorporation of a set of new slack-type decision variables,  $\tau_{i0}^-$ ,  $i = 1, \dots, m$ , and  $\tau_{r0}^+$ ,  $r = 1, \dots, s$ . The new formulation is enough to ensure that the projection points generated by the model are always Pareto-efficient with respect to the bounded technology  $T^{q_0}$ , something that cannot be claimed regarding the previous model (5). It is worth mentioning that Equation (8) is an extension of the weighted additive model in DEA (Lovell & Pastor 1995 or Cooper et al. 1999), which allows a quota system to be dealt with.

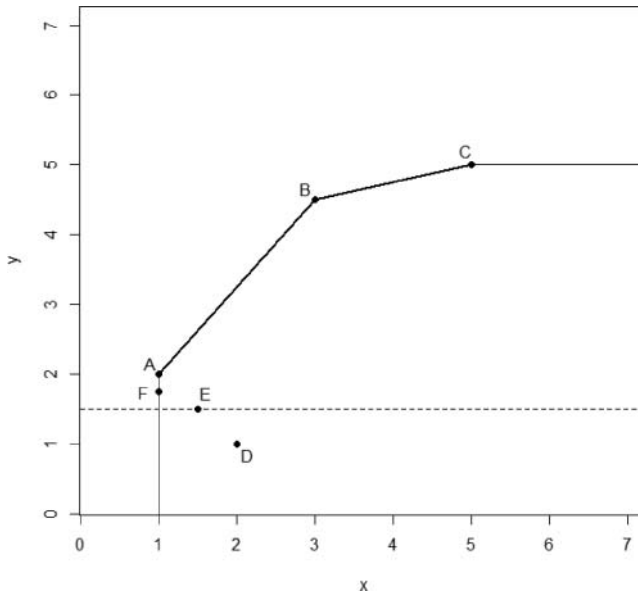
In the simple numerical example linked to Figure 1, D is projected onto point (1, 1.5), as desired, through model (8) thanks to the optimal solution:  $\lambda_{AD}^* = 1$ ,  $\lambda_{BD}^* = \lambda_{CD}^* = \lambda_{DD}^* = \lambda_{ED}^* = 0$ ,  $\tau_D^- = 0$ ,  $\tau_D^+ = 0.5$ ,  $s_D^- = 1$  and  $s_D^+ = 0.5$ . In fact, the obtained projection is the only strongly efficient point of  $T^{q_0}$ , as Figure 1 makes clear.

Although Equation (8) represents an extension of the weighted additive model, it has important changes with respect to the conventional model (2). The most obvious difference is, of course, the incorporation of a new constraint for limiting production (the restriction associated with the quota). Second, the projection points for the conventional additive model can be directly obtained from the optimal intensity variables as  $(\sum_{j=1}^n \lambda_{j0}^* x_{1j}, \dots, \sum_{j=1}^n \lambda_{j0}^* x_{mj}, \sum_{j=1}^n \lambda_{j0}^* y_{1j}, \dots, \sum_{j=1}^n \lambda_{j0}^* y_{sj})$ . However, in the case of Equation (8), new decision variables must be considered:  $(\sum_{j=1}^n \lambda_{j0}^* x_{1j} + \tau_{10}^{-*}, \dots, \sum_{j=1}^n \lambda_{j0}^* x_{mj} + \tau_{m0}^{-*}, \sum_{j=1}^n \lambda_{j0}^* y_{1j} - \tau_{10}^{+*}, \dots, \sum_{j=1}^n \lambda_{j0}^* y_{sj} - \tau_{s0}^{+*})$ . We are referring to  $\tau_{i0}^-$  for inputs and  $\tau_{r0}^+$  for outputs. Third, due to the addition of quota-related constraints, the optimal value of the conventional model is greater or equal than that associated with the new model.

**Proposition 3.1:** Let  $w_{i0}^- \in R_{++}^m$ ,  $i = 1, \dots, m$ , and  $w_{r0}^+ \in R_{++}^s$ ,  $r = 1, \dots, s$ . Then,  $\text{WA}(x_0, y_0, w^-, w^+) \geq \text{WA}^{q_0}(x_0, y_0, w^-, w^+)$ .

**Proof:** It is a consequence of the relationship  $T^{q_0} \subset T$ . ■





**Figure 2.** Illustration of dominated peers generated by model (8).

**3.2. The improvement of the basic model for production with quotas**

Let us now illustrate a situation that can be of interest from a managerial point of view. DEA models usually yield both efficiency values and benchmarking information for the evaluated units. This last information comes from the optimal value of the intensity variables ( $\lambda_{j0}^*$ ,  $j = 1, \dots, n$ ). A strictly positive value for this type of decision variables implies that  $DMU_j$  can be considered a peer or benchmark for  $DMU_0$ . At this point, we would like to explicitly make the distinction between projection point and peer in the sense that the projection generated by a DEA model is not necessarily an observation, whereas the peers are real observed units, which also play a role of benchmarks for the assessed DMUs. So, peers could be considered as relevant benchmarking information when the traditional DEA models are applied in practice, like for example, the weighted additive models (Lovell & Pastor 1995; Cooper et al. 1999, 2011). However, some peers determined by Equation (8) could be dominated in the sense of Pareto, something strange if we are seeking to satisfy Pareto-efficiency and the peers should make up the comparison group for the evaluated DMU. Figure 2 graphically illustrates this situation. Note that, in this figure, observation F has been incorporated to the above simple numerical example in two dimensions. In this case, if we evaluate unit D through model (8), we can get several optimal solutions, one of them is  $\lambda_{FD}^* = 1$ ,  $\lambda_{AD}^* = \lambda_{BD}^* = \lambda_{CD}^* = \lambda_{DD}^* = \lambda_{ED}^* = 0$ ,  $\tau_D^- = 0$ ,  $\tau_D^+ = 0.25$ ,  $s_D^- = 1$  and  $s_D^+ = 0.5$ . Unfortunately, unit F is not Pareto-efficient (see Figure 2), since it is dominated by unit A. So, it is possible that the manager of the assessed firm does not understand the comparison with a firm in the sample like F at all, since the performance of F may be enhanced. Clearly, we need to avoid these kinds of peers when model (8) is applied. In this sense, a question arises: Is it possible to reformulate model (8) in an equivalent way deleting units like F from the peers set? The answer is affirmative. Nevertheless, we first need to prove Proposition 3.2.

Before doing so, let us take a deeper insight into the structure of  $T^{q_0}$ , which will allow us to simplify the formulation of model (8) regarding the number of decision variables.

In this respect, let  $E$  be the subset of the initial sample of  $n$  DMUs that are Pareto-efficient for the unbounded technology,  $T$ . Then we get the following equivalence:

**Proposition 3.2:**  $T^{q_0}$  can be equivalently rewritten as

$$T^{q_0} = \{(x, y) \in R_+^{m+s} : (x, -y) \geq \sum_{j \in E} \lambda_j (x_j, -y_j), \sum_{j \in E} \lambda_j = 1, \lambda_j \geq 0, \forall j, y_1 \leq q_0\}. \quad (9)$$

**Proof:** On the one hand, let us denote by  $T_E^{q_0}$  the subset of  $R_+^m \times R_+^s$  given on the right-hand side of Equation (9). Since  $E \subset \{1, \dots, n\}$ , given a set of  $\lambda_j \geq 0, j \in E$ , with  $\sum_{j \in E} \lambda_j = 1$  such that  $(x, -y) \geq \sum_{j \in E} \lambda_j (x_j, -y_j)$ , we have that  $\bar{\lambda}_j \geq 0, j = 1, \dots, n$ , defined as  $\bar{\lambda}_j = \lambda_j, \forall j \in E$  and  $\bar{\lambda}_j = 0, \forall j \notin E$ , satisfy  $\sum_{j \in E} \lambda_j (x_j, -y_j) = \sum_{j=1}^n \bar{\lambda}_j (x_j, -y_j)$ . Consequently,  $T_E^{q_0} \subset T^{q_0}$ . On the other hand, we need to show that  $T_E^{q_0} \supset T^{q_0}$ . By Equation (9),  $\forall k = 1, \dots, n \exists \delta_j^k \geq 0, j \in E$ , with  $\sum_{j \in E} \delta_j^k = 1$  such that  $(x_k, -y_k) \geq \sum_{j \in E} \delta_j^k (x_j, -y_j)$  since each DMU $_k$  belongs to  $T$ . Now, for any  $(x, y) \in T^{q_0}$  we have that, by Equation (4),  $\exists \bar{\lambda}_z \geq 0, z = 1, \dots, n$ , with  $\sum_{z=1}^n \bar{\lambda}_z = 1$  such that  $(x, -y) \geq \sum_{z=1}^n \bar{\lambda}_z (x_z, -y_z)$  and  $y_1 \leq q_0$ . Note then that  $(x, -y) \geq \sum_{z=1}^n \bar{\lambda}_z (x_z, -y_z) \geq \sum_{z=1}^n \bar{\lambda}_z \sum_{j \in E} \delta_j^z (x_j, -y_j) = \sum_{j \in E} [\sum_{z=1}^n \bar{\lambda}_z \delta_j^z] (x_j, -y_j)$ . Now, defining  $\lambda_j := \sum_{z=1}^n \bar{\lambda}_z \delta_j^z, j \in E$ , we can check that  $\lambda_j = \sum_{z=1}^n \bar{\lambda}_z \delta_j^z \geq 0$  for all  $j \in E$  and that, additionally,  $\sum_{j \in E} \lambda_j = \sum_{j \in E} \sum_{z=1}^n \bar{\lambda}_z \delta_j^z = \sum_{z=1}^n \sum_{j \in E} \bar{\lambda}_z \delta_j^z = \sum_{z=1}^n \bar{\lambda}_z \sum_{j \in E} \delta_j^z = \sum_{z=1}^n \bar{\lambda}_z = 1$ . In this way,  $(x, y) \in T_E^{q_0}$  and, therefore,  $T_E^{q_0} \supset T^{q_0}$ . ■

According to Proposition 3.2, we can refine model (8) by substituting ‘ $j = 1, \dots, n$ ’ by ‘ $j \in E$ ’ in the sums that appear in the constraints of the model. Hence, the efficient projection of  $(x_0, y_0), (x_0^*, y_0^*) = (\sum_{j \in E} \lambda_{j0}^* x_j + \tau_0^{*-}, \sum_{j \in E} \lambda_{j0}^* y_j - \tau_0^{*+})$ , will only ever consider convex combinations of points of  $E$  together with a set of slacks. So, model (8) can be equivalently written as Equation (10).

$$\begin{aligned} \text{WA}^{q_0}(x_0, y_0, w^-, w^+) = \text{Max} \quad & \sum_{i=1}^m w_{i0}^- s_{i0}^- + \sum_{r=1}^s w_{r0}^+ s_{r0}^+ \\ \text{s.t.} \quad & \sum_{j \in E} \lambda_{j0} x_{ij} + \tau_{i0}^- = x_{i0} - s_{i0}^-, \quad i = 1, \dots, m \quad (10.1) \\ & \sum_{j \in E} \lambda_{j0} y_{rj} - \tau_{r0}^+ = y_{r0} + s_{r0}^+, \quad r = 1, \dots, s \quad (10.2) \\ & \sum_{j \in E} \lambda_{j0} y_{1j} - \tau_{10}^+ \leq q_0, \quad (10.3) \\ & \sum_{j \in E} \lambda_{j0} = 1, \quad (10.4) \\ & \lambda_{j0} \geq 0, \quad j \in E \quad (10.5) \\ & s_{i0}^- \geq 0, \quad i = 1, \dots, m \quad (10.6) \\ & s_{r0}^+ \geq 0, \quad r = 1, \dots, s \quad (10.7) \\ & \tau_{i0}^- \geq 0, \quad i = 1, \dots, m \quad (10.8) \\ & \tau_{r0}^+ \geq 0, \quad r = 1, \dots, s \quad (10.9) \end{aligned}$$

The use of model (10) needs a simple pre-processing stage to identify the subset  $E$  of Pareto-efficient DMUs in the sample for the ‘unbounded’ production possibility set  $T$ . The procedure consists in running the ‘traditional’ additive model (2) with  $w_{i0}^- = 1$ ,  $i = 1, \dots, m$ , and  $w_{r0}^+ = 1$ ,  $r = 1, \dots, s$ , for each observation and checking the optimal value. Only an optimal value of zero signals that we are facing a Pareto-efficient unit.

In this way, model (10) presents at least two interesting advantages in comparison with model (8). First, it prevents dominated units like F appearing as peers of some evaluated observations. We can ensure that all the peers will be Pareto-efficient DMUs by applying Equation (10) instead of the original model (8). Second, from a computational point of view, model (10) has a less number of decision variables than model (8). In particular, it presents a saving of  $n - \text{card}(E)$  decision variables in the model, where  $\text{card}(E)$  denotes the cardinal of set  $E$ , although, of course, the application of Equation (10) requires a pre-processing step that Equation (8) does not need.

We now turn to the characterization of the Pareto-efficient frontier of  $T^{q_0}$ . For the unbounded technology  $T$ , it is well known that a DMU is Pareto-efficient if and only if the optimal value of the weighted additive model is zero (see Proposition 2.1). Regarding technology based on a quota regime  $T^{q_0}$ , we are also able to prove an analogous result: a DMU is Pareto-efficient if and only if the optimal value of model (10) is zero. This result can be seen as the counterpart of Proposition 2.1 for the weighted additive model in DEA in the context of production under output quotas.

**Proposition 3.3:**  $(x_0, y_0)$  is a Pareto-efficient point of  $T^{q_0}$  if and only if the optimal value of model (10) is zero.

**Proof:** Let  $(\lambda_0^*, s_0^{-*}, s_0^{+*}, \tau_0^{-*}, \tau_0^{+*})$  be an optimal solution of model (10). We first prove that if  $(x_0, y_0)$  is a Pareto-efficient point of  $T^{q_0}$ , then  $\sum_{i=1}^m w_{i0}^- s_{i0}^{-*} + \sum_{r=1}^s w_{r0}^+ s_{r0}^{+*} = 0$ . To this end, let us assume that this claim is false and we will arrive at a contradiction. If  $\sum_{i=1}^m w_{i0}^- s_{i0}^{-*} + \sum_{r=1}^s w_{r0}^+ s_{r0}^{+*} > 0$ , then, by Equations (10.1) and (10.2), we have that  $x_{i0}^* := \sum_{j \in E} \lambda_{j0}^* x_{ij} + \tau_{i0}^{-*} \leq x_{i0}$ ,  $\forall i = 1, \dots, m$ ,  $y_{r0}^* := \sum_{j \in E} \lambda_{j0}^* y_{rj} - \tau_{r0}^{+*} \geq y_{r0}$ ,  $\forall r = 1, \dots, s$ , and such that, at least for one input  $i'$  or one output  $r'$ , the previous inequalities are strict since  $s_{i0}^{-*}, \forall i, s_{r0}^{+*}, \forall r$ , cannot be all zero. Additionally, by Proposition 3.2,  $(x_0^*, y_0^*) \in T^{q_0}$  since  $\tau_{i0}^{-*} \geq 0, \forall i$ , and  $\tau_{r0}^{+*} \geq 0, \forall r$ . Finally, point  $(x_0, y_0)$  is strictly dominated by point  $(x_0^*, y_0^*)$ , thus leading to a contradiction. Let us now prove that, given  $(\lambda_0^*, s_0^{-*}, s_0^{+*}, \tau_0^{-*}, \tau_0^{+*})$ , if  $\sum_{i=1}^m w_{i0}^- s_{i0}^{-*} + \sum_{r=1}^s w_{r0}^+ s_{r0}^{+*} = 0$ , then  $(x_0, y_0)$  is a Pareto-efficient point of  $T^{q_0}$ . In order to do this, let us assume that there exists a point  $(\tilde{x}, \tilde{y}) \in T^{q_0}$  with  $(\tilde{x}, \tilde{y}) \neq (x_0, y_0)$  such that  $\tilde{x}_i \leq x_{i0}, \forall i = 1, \dots, m$ , and  $\tilde{y}_r \geq y_{r0}, \forall r = 1, \dots, s$ . Then, by Proposition 3.2,  $\exists \tilde{\lambda}_j \geq 0, j \in E$ , with  $\sum_{j \in E} \tilde{\lambda}_j = 1$  such that  $\tilde{x}_i \geq \sum_{j \in E} \tilde{\lambda}_j x_{ij}, i = 1, \dots, m$ ,  $\tilde{y}_r \leq \sum_{j \in E} \tilde{\lambda}_j y_{rj}, r = 1, \dots, s$ , and  $\tilde{y}_1 \leq q_0$ . Now we can define  $\tilde{\tau}_i^- := \tilde{x}_i - \sum_{j \in E} \tilde{\lambda}_j x_{ij}, \tilde{s}_i^- := x_{i0} - \tilde{x}_i, i = 1, \dots, m$ , and  $\tilde{\tau}_r^+ := \sum_{j \in E} \tilde{\lambda}_j y_{rj} - \tilde{y}_r, \tilde{s}_r^+ := \tilde{y}_r - y_{r0}$ . Then, it is easy to check that  $(\tilde{\lambda}, \tilde{s}^-, \tilde{s}^+, \tilde{\tau}^-, \tilde{\tau}^+)$  is a feasible solution of model (10). Also, some  $\tilde{s}_i^-, i = 1, \dots, m, \tilde{s}_r^+, r = 1, \dots, s$ , is strictly positive due to our hypothesis on  $(\tilde{x}, \tilde{y})$ . Consequently,  $\sum_{i=1}^m w_{i0}^- \tilde{s}_i^- + \sum_{r=1}^s w_{r0}^+ \tilde{s}_r^+ > \sum_{i=1}^m w_{i0}^- s_{i0}^{-*} + \sum_{r=1}^s w_{r0}^+ s_{r0}^{+*} = 0$ , which is a contradiction. ■

Next, we prove the counterpart of Proposition 2.2 in the context of producing with output bounds. Proposition 3.4 establishes that the target point determined by programme (10) is always on the strongly efficient frontier of the bounded technology.

**Proposition 3.4:** Let  $w_{i0}^- \in R_{++}^m$ ,  $i = 1, \dots, m$ , and  $w_{r0}^+ \in R_{++}^s$ ,  $r = 1, \dots, s$ , and let  $(\lambda_0^*, s_0^{-*}, s_0^{+*}, \tau_0^{-*}, \tau_0^{+*})$  be an optimal solution of Equation (10). Then  $(x_0^*, y_0^*) := (\sum_{j \in E} \lambda_{j0}^* x_j + \tau_0^{-*}, \sum_{j \in E} \lambda_{j0}^* y_j - \tau_0^{+*})$  belongs to  $\partial^s(T^{q_0})$ .

**Proof:** Let  $(\lambda_0^*, s_0^{-*}, s_0^{+*}, \tau_0^{-*}, \tau_0^{+*})$  be an optimal solution of model (10). Then, it is obvious that  $(x_0^*, y_0^*) \in T^{q_0}$ . Let us assume that there exists a point  $(\tilde{x}, \tilde{y}) \in T^{q_0}$  with  $(\tilde{x}, \tilde{y}) \neq (x_0^*, y_0^*)$  such that  $\tilde{x}_i \leq x_{i0}^*$ ,  $\forall i = 1, \dots, m$ , and  $\tilde{y}_r \geq y_{r0}^*$ ,  $\forall r = 1, \dots, s$ . Then, by Proposition 3.2,  $\exists \tilde{\lambda}_j \geq 0$ ,  $j \in E$ , with  $\sum_{j \in E} \tilde{\lambda}_j = 1$  such that  $\tilde{x}_i \geq \sum_{j \in E} \tilde{\lambda}_j x_{ij}$ ,  $i = 1, \dots, m$ ,  $\tilde{y}_r \leq \sum_{j \in E} \tilde{\lambda}_j y_{rj}$ ,  $r = 1, \dots, s$ , and  $\tilde{y}_1 \leq q_0$ . Now we can define  $\tilde{\tau}_i^- := \tilde{x}_i - \sum_{j \in E} \tilde{\lambda}_j x_{ij}$ ,  $\tilde{s}_i^- := x_{i0} - (\sum_{j \in E} \tilde{\lambda}_j x_{ij} + \tilde{\tau}_i^-) = x_{i0} - \tilde{x}_i \geq x_{i0} - x_{i0}^* = s_{i0}^{-*} \geq 0$ ,  $i = 1, \dots, m$ , and  $\tilde{\tau}_r^+ := \sum_{j \in E} \tilde{\lambda}_j y_{rj} - \tilde{y}_r$ ,  $\tilde{s}_r^+ := (\sum_{j \in E} \tilde{\lambda}_j y_{rj} - \tilde{\tau}_r^+) - y_{r0} = \tilde{y}_r - y_{r0} \geq y_{r0}^* - y_{r0} = s_{r0}^{+*} \geq 0$ ,  $r = 1, \dots, s$ . Then, it is easy to check that  $(\tilde{\lambda}, \tilde{s}^-, \tilde{s}^+, \tilde{\tau}^-, \tilde{\tau}^+)$  is a feasible solution of model (10). Also, regarding the objective function,  $\sum_{i=1}^m w_{i0}^- \tilde{s}_i^- + \sum_{r=1}^s w_{r0}^+ \tilde{s}_r^+ > \sum_{i=1}^m w_{i0}^- s_{i0}^{-*} + \sum_{r=1}^s w_{r0}^+ s_{r0}^{+*}$  since  $(\tilde{x}, \tilde{y}) \neq (x_0^*, y_0^*)$ . This last result is in contradiction to the fact that  $(\lambda_0^*, s_0^{-*}, s_0^{+*}, \tau_0^{-*}, \tau_0^{+*})$  is an optimal solution of model (10). ■

In the next section, we show how model (10) performs in comparison with the unbounded framework, resorting to a real data-set consisting of Canadian provinces and milk production. Additionally, we show that the decision variable  $\tau_0^+$  is needed in order to correctly measure technical inefficiency under an additive-type approach.

## 4. An application of the new approach: milk production under a quota regime in Canada

### 4.1. Data-set

This study uses data from Canadian dairy farms at the level of 10 Canadian provinces in the year 2009. According to the Canadian Dairy Commission in 2015 about 82% of Canada's dairy farms were in Ontario and Quebec, 13% in the Western provinces (that is in Manitoba, Saskatchewan, Alberta and British Columbia) and 5% in Atlantic provinces (that is in New Brunswick, Prince Edward Island, Nova Scotia and Newfoundland and Labrador). The data applied in this study are provided by two sources: the Canadian Dairy Information Centre and Statistics Canada.

Dairy farms combine the production of milk with the production of meat (beef and veal). Therefore, the production model is specified in two outputs: milk production at the farm (expressed in hectolitres) and cattle revenues (in Canadian dollars). The first variable was supported by the Canadian Dairy Information Centre, while the second by Statistics Canada. Milk output is restricted by the quota system and the values of quotas (in hectolitres) were obtained from the Canadian Dairy Information Centre. It is worth noticing that the data on quota is reported in million kg of butterfat, and therefore it was necessary to recalculate and convert it into hectolitres.

**Table 1.** Descriptive statistics of input/output data and quota for Canadian provinces, year 2009.

Variable	Mean	Standard deviation	Minimum	Maximum
<b>Outputs</b>				
<i>Milk production</i>				
All provinces	7,662,781.7	10,306,796.7	484,092.0	28,457,600.0
Quebec and Ontario	26,740,131.5	2,428,867.2	25,022,663.0	28,457,600.0
Western provinces	4,644,078.5	2,233,641.8	2,307,076.0	6,609,935.0
Atlantic provinces	1,142,810.0	523,676.1	484,092.0	1,714,545.0
<i>Cattle revenues</i>				
All provinces	55,097.0	57,913.3	15,343.0	204,636.0
Quebec and Ontario	20,174.5	6,832.8	15,343.0	25,006.0
Western provinces	100,617.5	72,589.2	40,977.0	204,636.0
Atlantic provinces	27,037.8	12,307.6	15,438.0	44,250.0
<b>Inputs</b>				
<i>Number of cows</i>				
All provinces	96.4	130.9	6.4	361.9
Quebec and Ontario	339.7	31.5	317.4	361.9
Western provinces	56.2	24.8	28.2	80.7
Atlantic provinces	15.1	6.9	6.4	21.8
<i>Total costs</i>				
All provinces	618,874.5	342,626.2	315,922.0	1,391,508.0
Quebec and Ontario	336,126.0	28,572.8	315,922.0	356,330.0
Western provinces	733,229.8	176,710.8	517,140.0	949,252.0
Atlantic provinces	645,893.5	498,970.9	347,893.0	1,391,508.0
<b>Quota</b>				
All provinces	8,264,241.0	11,294,166.4	507,841.1	31,440,124.4
Quebec and Ontario	29,179,820.4	3,196,552.6	26,919,516.4	31,440,124.4
Western provinces	4,863,155.9	2,344,636.0	2,387,894.4	6,881,523.6
Atlantic provinces	1,207,536.4	557,702.2	507,841.1	1,823,487.5

Note: Milk production and quota are expressed in hectolitres. The number of cows is in thousands. Cattle revenues and total costs are in Canadian dollars (CAD). Western provinces – Manitoba, Saskatchewan, Alberta and British Columbia; Atlantic provinces – New Brunswick, Prince Edward Island, Nova Scotia and Newfoundland and Labrador.

There is a variety of inputs that could be considered to model dairy production that were used in previous studies, such as number of cows, labour, fixed assets, agricultural area, material and energy-related costs, and expenditure on feed (Mbaga et al. 2003; Emvalomatis et al. 2011). Due to the availability of data, as well as to obtain certain accuracy in our estimations (using too many inputs could render too many efficient observations given a relatively small sample size), two main inputs are considered in the application: number of cows (in thousands) and total operating costs (in Canadian dollars), both taken from Statistics Canada. Total aggregate operating costs, among others, feed expenses, machinery expenses and salaries.

Summary statistics of input–output variables and quota for all provinces and per province group (Quebec and Ontario, Western provinces and Atlantic provinces) in 2009 appear in Table 1. These data show some variability in the sample, as shown by the values of standard deviations relative to their respective means. The highest variation is exhibited by milk production, number of cows and quota. Furthermore, it is worth noticing that the province of Newfoundland and Labrador is the one that exhibits the smallest values in the number of cows, milk production and consequently the quota; however, it incurs the largest total costs. The lowest values of total costs are incurred in Quebec, but at the same time, this province has the highest values in the number of cows, milk production and quota. The second output, cattle revenues, is the largest for Alberta, and the lowest for Quebec. Hence, the data show that in 2009, Quebec is the Canadian province with the largest number of dairy farms and produces the largest amount of milk, but at the same time it seems to neglect the second output of cattle. On the contrary, the Atlantic province

of Newfoundland and Labrador is characterized by the smallest milk production, while the Western province of Alberta seems to focus on the second output of cattle, reaching the largest values in cattle revenues. These observations are further strengthened by the results of descriptive statistics per type of province which show that Quebec and Ontario, on average, have the largest values in milk production, but the smallest of cattle revenues, while Atlantic provinces produce the lowest values of milk and Western provinces obtain the largest values of revenues from cattle, on average.

## 4.2. Results

In the empirical application of model (10) with quota and the unbounded model without the quota restriction (2), the following weights were applied:  $w_{i0}^- = \frac{1}{x_{i0}}$ ,  $i = 1, \dots, m$  (inputs),  $w_{r0}^+ = \frac{1}{y_{r0}}$ ,  $r = 1, \dots, s$  (outputs). The usage of such weights yields that the models coincide with the Measure of Inefficiency Proportions (MIP) (see Cooper et al. 1999) that measures the aggregation of inefficiency proportions due to the excessive input usage and shortfalls in output production. Worthy of note is that MIP is not bounded by 1 as in radial inefficiency measures. The full efficiency is attained with  $MIP = 0$ .

Table 2 reports the inefficiency results of the model that takes into account quota, and the unbounded model without quota considerations for each of the Canadian provinces separately and on average for the year 2009. The table also summarizes the results for each province group (Quebec and Ontario, Western provinces and Atlantic provinces). In addition, the table also contains the results of slack-type decision variable  $\tau_{r0}^+$  for the output for which a quota is imposed, that is, for milk production.

Several observations can be made based on the results in Table 2. First, on average, ‘inefficient’ Canadian provinces with regard to milk production are (slightly) more inefficient when quota is not taken into account than when the inefficiency model considers quota restriction (inefficiency value of 1.0629 versus 1.0388, respectively). When considering all provinces in the sample, the inefficiency without quota is equal to 0.4252, while subject to quota it equals 0.4155. Hence, more inefficiency is found in the model without

**Table 2.** Inefficiency results for the model, with and without quota restrictions for Canadian provinces, year 2009.

Province	Inefficiency with quota	Inefficiency without quota	$\tau_{r0}^+$ for output with quota
British Columbia	0.0000	0.0000	0.0000
Alberta	0.0000	0.0000	0.0000
Saskatchewan	0.1174	0.1174	0.0000
Manitoba	1.5720	1.6051	$1.0460 \times 10^5$
Ontario	0.0000	0.0000	0.0000
Quebec	0.0000	0.0000	0.0000
New Brunswick	1.5781	1.6417	$0.8630 \times 10^5$
Nova Scotia	0.8876	0.8876	0.0000
Prince Edward Island	0.0000	0.0000	0.0000
Newfoundland and Labrador	0.0000	0.0000	0.0000
Mean (all units)	0.4155	0.4252	$0.1909 \times 10^5$
Mean (inefficient units)	1.0388	1.0629	$0.9545 \times 10^5$
Ontario and Quebec	0.0000	0.0000	0.0000
Western provinces	0.4224	0.4306	$0.2620 \times 10^5$
Atlantic provinces	0.6164	0.6323	$0.2160 \times 10^5$

Note: Western provinces – Manitoba, Saskatchewan, Alberta and British Columbia; Atlantic provinces – New Brunswick, Prince Edward Island, Nova Scotia and Newfoundland and Labrador.

**Table 3.** Targets and peers for the model with quota restrictions for Canadian provinces, year 2009.

Province	Targets				Peers
	Outputs		Inputs		
	Milk production	Cattle revenues	Number of cows	Total expenses	
(1) British Columbia	6,609,935.0	93,362.0	72.9	949,252.0	1
(2) Alberta	6,498,846.0	204,636.0	80.7	720,620.0	2
(3) Saskatchewan	2,307,095.8	63,495.1	28.2	658,331.5	1, 2, 9, 10
	(0.00%)	(0.00%)	(0.00%)	(-11.74%)	
(4) Manitoba	3,332,791.2	102,799.7	42.9	512,507.9	2, 9
	(5.45%)	(150.87%)	(0.00%)	(-0.90%)	
(5) Ontario	25,022,663.0	25,006.0	317.4	356,330.0	5
(6) Quebec	28,457,600.0	15,343.0	361.9	315,922.0	6
(7) New Brunswick	1,423,528.8	38,681.2	19.1	381,475.7	2, 9
	(4.92%)	(150.56%)	(0.00%)	(-2.32%)	
(8) Nova Scotia	1,728,547.2	45,956.6	21.8	396,343.8	2, 9
	(0.82%)	(75.31%)	(0.00%)	(-12.63%)	
(9) Prince Edward Island	1,015,807.0	22,248.0	13.0	347,893.0	9
(10) Newfoundland and Labrador	484,092.0	44,250.0	6.4	1,391,508.0	10

quota consideration. From a conceptual model point of view, this result may not seem so surprising since in the model with quota, the bounds are set to the technology that results in units being closer to the frontier (that is, being more efficient) than in the model without quota consideration. From an empirical perspective, this result is significant since it might imply that quota imposition by a market regulator can render greater efficiency in firms' production processes. That finding can imply that quota exchanges ensure that quota transfers to the most efficient producers. Hence, the arguments that criticize quota system claiming that such a system shields producers from competition, can be confronted with the results of our study that provide arguments for supporters of the quota system. That finding can have implications for policy-makers in defining policy instruments that enhance the efficiency of economic sectors, providing arguments in favour of quota system.

Second, looking at the results for individual provinces, it can be concluded that the differences in average results for both models originate in different inefficiency outcomes for two provinces: Manitoba and New Brunswick. In particular, Manitoba is found to be more efficient by 2% and New Brunswick by 4% in the model with quota than that without quota. The inefficiency of Manitoba at 1.5720 in the model with quota implies that dairy farms in this province should reduce, on average, the usage of inputs by 39% and increase the production of outputs by 39% [ $1.5720$  (inefficiency)/4 (number of inputs and outputs in the model)].

Third, the results per province group show that the provinces with the largest number of dairy farms in Canada (Ontario and Quebec) are also the provinces that are efficient (that is have inefficiency values equal to 0) according to both models, with and without quota consideration. Western provinces, which are the second largest in terms of the size of dairy industry in Canada, also hit the second with regard to their average inefficiency values (0.4224 in the model with quota and 0.4306 in the model without quota). Finally, Atlantic provinces, which represent the smallest share of the dairy industry in Canada, are found to be the most inefficient group with an average inefficiency of 0.6164 and 0.6323 in the model with quota and without quota, respectively.

Fourth, the results of the slack-type decision variable  $\tau_{r_0}^+$  for the output of milk production reveal that precisely for these provinces for which the differences in models with quota and without quota are observed (Manitoba and New Brunswick), this variable has values different from 0. This finding shows, in the context of the weighted additive model, the relevance of using the new model (10) with the slack-type decision variables rather than model (5).

We further analyse the empirical results for the model with quota more thoroughly. Table 3 shows the targets and peers for each of the Canadian provinces for the year 2009 with regard to this model. Targets represent the target values for inputs and outputs (which for efficient provinces obviously correspond to actual values for inputs used and outputs produced by these provinces), while peers indicate the provinces that should serve as benchmarks. As for the targets for inefficient provinces, the percentages of improvement in each variable that are needed to achieve the efficiency are also reported between parentheses.

The first result to note from Table 3 is that the efficient provinces of Alberta and Prince Edward Island serve as a reference benchmark for the greatest number of provinces. Turning to the results for targets, the upshot of the table indicates that in order to reach efficiency, the province of Saskatchewan should focus exclusively on the reduction of the input of total costs through a decrease in its value by almost 12%. The results also suggest the need for substantial improvements in almost all variables for the provinces of Manitoba and New Brunswick, which were found to be the most inefficient. It can also be appreciated from the table that the output of cattle revenues is the variable that requires the largest improvements across all inefficient provinces. On the contrary, the results show that the input of number of cows does not require any improvement at all. It is important to remember that the application of model (10) guarantees that computed targets are always on the strongly efficient frontier and that yielded peers are always non-dominated in the sense of Pareto.

## 5. Conclusions

In this paper, an extension of the weighted additive model in DEA has been introduced for evaluating technical inefficiency of units that operate in regulated markets under production quotas. The adaptation of the traditional weighted additive model to these types of situations can be useful, from the point of view of practice, to managers and policy-makers in their decision-making, by providing both performance and benchmarking information, while ensuring the measurement of technical efficiency satisfying Pareto-efficiency.

The model of production under a quota regime has been empirically applied in this paper to determine the technical inefficiency of Canadian provinces with regard to milk produced by dairy farms in the year 2009, and it has been compared with a conventional model in which output is not bounded by any quota. The results have shown smaller inefficiency outcomes in Canadian provinces under models governed by quotas than those without quota impositions, on average. The bounds of technology due to quota can therefore have some implications for policy-makers who plan interventions in economic sectors in an attempt to improve their performance and competitiveness. Nevertheless, this paper has shown only slight differences in inefficiency between models with and without quota, and future research studies could analyse these issues using larger data-sets and other contexts to possibly depict more significant heterogeneities in results.



## Acknowledgments

We would like to express our gratitude to two anonymous referees for their helpful comments, as well as to the Spanish Ministry for Economy and Competitiveness for supporting us through grant MTM2013-43903-P.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

Spanish Ministry for Economy and Competitiveness [grant number MTM2013-43903-P].

## References

- Banker RD, Charnes A, Cooper WW. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage Sci.* 30:1078–1092.
- Bunting J. 2009. Canadian farm milk quota system yields rewards to producers, rural dairy communities. *Milkweed.* 355:8–9.
- Canadian Dairy Commission [Internet]. Ottawa: Canadian government; [cited 2016 Aug 10]. Available from: <http://www.cdc-ccl.gc.ca/CDC/index-eng.php>
- Canadian Dairy Information Centre [Internet]. Ottawa: Canadian government; [cited 2016 July 20]. Available from: [http://www.dairyinfo.gc.ca/index\\_e.php](http://www.dairyinfo.gc.ca/index_e.php)
- Charnes A, Cooper WW, Golany B, Seiford L, Stutz J. 1985. Foundations of data envelopment analysis for pareto-koopmans efficient empirical production functions. *J Econ.* 30:91–107.
- Charnes A, Cooper WW, Rhodes E. 1978. Measuring the efficiency of decision making units. *Eur J Oper Res.* 2:429–444.
- Cooper WW, Park KS, Pastor JT. 1999. RAM: a range adjusted measure of inefficiency for use with additive models, and relations to others models and measures in DEA. *J Prod Anal.* 11:5–42.
- Cooper WW, Pastor JT, Borrás F, Aparicio J, Pastor D. 2011. BAM: a bounded adjusted measure of efficiency for use with bounded additive models. *J Prod Anal.* 35:85–94.
- Emvalomatis G, Stefanou SE, Oude Lansink A. 2011. A reduced-form model for dynamic efficiency measurement: application to dairy farms in germany and the netherlands. *Am J Agric Econ.* 93:161–174.
- Färe R, Grosskopf S, Lovell CAK. 1985. *The measurement of efficiency of production.* Dordrecht (NL): Springer.
- Farrell MJ. 1957. The measurement of productive efficiency. *J R Stat Soc.* 120:253–281.
- Koopmans TC. 1951. Analysis of production as an efficient combination of activities. In: Koopmans TC, editor. *Activity analysis of production and allocation.* New York (NY): Wiley; p. 33–97.
- Lovell CAK, Pastor JT. 1995. Units invariant and translation invariant DEA models. *Oper Res Lett.* 18:147–151.
- Mbaga MD, Romain R, Larue B, Lebel L. 2003. Assessing technical efficiency of Québec dairy farms. *Can J Agric Econ.* 51:121–137.
- Pastor JT, Ruiz JL, Sirvent I. (1999). An enhanced DEA Russell graph efficiency measure. *Eur J Oper Res.* 115:596–607.
- Statistics Canada [Internet]. Ottawa: Canadian Government; [cited 2016 July 20]. Available from: <http://www.statcan.gc.ca/eng/>
- Tone K. 2001. A slacks-based measure of efficiency in data envelopment analysis. *Eur J Oper Res.* 130:498–509.