NJAS - Wageningen Journal of Life Sciences 70–71 (2014) 33–38

Contents lists available at ScienceDirect

NJAS - Wageningen Journal of Life Sciences

journal homepage: www.elsevier.com/locate/njas

Differences in TFP growth among groups of dairy farms in the Netherlands

T.H. Keizer [a,*], G. Emvalomatis [b]

[a] Business Economics Group, Wageningen University, 6707 KN Wageningen, The Netherlands
[b] Economic Studies, University of Dundee, DD1 4HN, Dundee, UK

ARTICLE INFO

Article history:
Received 27 June 2012
Received in revised form 21 January 2013
Accepted 4 March 2014
Available online 12 April 2014

Keywords:
dairy farming
productivity growth
farm size
intensity

ABSTRACT

The trend towards fewer and larger farms characterising agriculture in most industrialised countries can be partly attributed to larger farms becoming more productive by exploiting the economies of scale inherent in the production technology they employ. This paper examines whether larger and more intensive dairy farms in the Netherlands have been experiencing faster productivity growth than smaller farms, with the objective of determining which types of farms are more likely to prosper in the long run. Classification and regression trees are proposed as a valid way of classifying farms according to their size and farming intensity. At a second stage total factor productivity growth is calculated and decomposed into technical progress, efficiency change and scale effects for each class of farms using Data Envelopment Analysis. The results suggest that, for all classes of farms, productivity growth is driven almost exclusively by technical progress. The rate of technical progress has been higher for large intensive farms, implying that recent technical innovations are more beneficial to this type of dairy farms.

1. Introduction

The dairy sector in the Netherlands, as in most industrialized countries, has undergone major structural changes over the past decades. First, while half a century ago most farms produced an array of outputs, today they are highly specialised in the production of milk, with the bulk of grain feed bought in the market. Second, the number of dairy farms has been steadily declining, while the size of the farms that remain operational has been increasing. From an economic point of view, larger farms are possibly becoming more productive by exploiting the economies of scale inherent in the production technology. This, in turn, makes larger farms more competitive relative to the smaller ones, as they can produce more output per unit of input. Although the Netherlands is a country with a uniform climate, today's Dutch dairy farms are quite heterogeneous in terms of size. The 5% smallest dairy farms have less than 20 cows and have a gross margin as low as €54,000, whereas the 5% largest farms have more than 150 cows and their gross margin is larger than €318,000. From these differences in such a small country, a question rises naturally: are the trends towards fewer and larger farms going to persist? Furthermore, are benefits arising from economies of scale the only reason behind these trends?

The role of productivity growth is of major importance in answering these questions. Newman and Matthews [1] argue that differences in productivity growth rates is the main reason behind divergent trends in competitiveness. That is, if larger farms consistently experience faster productivity growth then, in the long run, they will become more competitive and, consequently, encourage smaller farms to adjust by expanding their scale or be driven out of business, with larger farms possibly acquiring their assets. Therefore, the evolution of productivity over time provides an indication of the development of relative competitiveness of different farming systems.

Before proceeding with the analysis we need an operational definition of productivity. In farming systems where individual farms use multiple inputs to produce multiple outputs, the definition of productivity is not straightforward. Total Factor Productivity (TFP) growth, defined as growth in outputs that cannot be attributed to growth in inputs, is the most widely accepted measure of productivity growth in such a setting. Apart from its ability to accommodate technologies where multiple inputs are used and multiple outputs produced, TFP growth encompasses the dynamic nature of the questions considered in this article.

TFP growth can be measured using different methods, each one of which has different data requirements and relies on alternative

* Corresponding author.
E-mail addresses: Teunkeizer@gmail.com (T.H. Keizer), Gemvalom@gmail.com (G. Emvalomatis).

http://dx.doi.org/10.1016/j.njas.2014.03.001
1573-5214/

assumptions about the representation of the production technology and the ability of the farm manager to exploit the full potential of the technology. Simple Törnqvist indexes can be constructed using data on each farm in isolation of other farms in the sample, but under the assumption that every farm is technically efficient [2]. Two methods that explicitly recognize that farms may be inefficient in transforming inputs into outputs are Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA). Since the introduction of SFA by Aigner *et al.* [3] and Meeusen and Broeck [4] this method has been widely applied. The other method was first proposed by Farrell [5] and extended by Seitz [6]. It presented a way of estimating a frontier production using linear programming techniques, but did not get wide acceptance until Charnes *et al.* [2] formalized it and coined the term Data Envelopment Analysis.

Both DEA and SFA compare multiple inputs with multiple outputs to measure efficiency and productivity and they can decompose TFP growth into three components: technical change, technical efficiency change and scale effect. Technical change results from a shift in the production technology. Technical efficiency comes from the farm's ability to use the available technology without wasting resources and the ability of a farm to use its inputs more efficiently and by operating closer to the technology frontier [7]. The last component, the scale effect, captures the effect on productivity of the ability of the farm to exploit economies of scale by modifying its size. The DEA method involves the use of linear programming methods to construct a non-parametric frontier over the data. Efficiency scores are then calculated relative to this frontier [7]. This frontier is not approximated by a production function as in the SFA method, but it is formed by dairy farms which produce the maximum possible amount of output with a given amount of inputs. Inefficient farms are projected onto the production frontier by using a convex combination of efficient farms (peers) that use similar input and output mixes [8].

Both methods available for calculating and decomposing TFP growth implicitly assume that all dairy farms share the same production technology. In our case, however, different types of farms could be employing alternative technologies, which are precisely the ones that are best-suited for them, given their size or other farm characteristics. Calculating TFP growth for all dairy farms would assume a common frontier, but if different types of farms employ alternative technologies this approach is not appropriate. The latent-class stochastic frontier approach [9] has been proposed in the literature as a way of dealing with this issue. However, the latent-class approach can be applied only in a parametric setting, i.e. when using SFA. Additionally, because the number of parameters to be estimated is a multiple of the number of assumed technologies, the latent-class framework becomes impractical in applications when the number of inputs and outputs considered is large. In this article we use instead a classification and regression tree (CRT) to distinguish classes. CRT's can divide the dairy farms in groups based on the technologies they employ and TFP growth can be calculated for each group of dairy farms which share the same technology using non-parametric techniques.

The aim of this study is to determine which types of farms are more likely to prosper in the long run based on the calculation of TFP growth. To achieve this objective the dairy farms are separated in groups in terms of the production technology they employ. The average TFP growth rate for each of these groups is calculated using DEA.

The rest of the paper is organised as follows. The next section presents the methodology used in this article. Section 3 contains a description of the data and section 4 presents and discusses the empirical findings. Finally, section 5 provides some concluding comments.

2. Methodology

Classification and Regression Trees (CRT) is a technique which can divide the farms in groups based on their characteristics. CRT splits the data into segments that are as homogeneous as possible with respect to the dependent variable. The most important difference between a classification tree and a regression tree is that classification trees use discrete and categorical dependent variables, whereas the dependent variable in regression trees is continuous. A CRT is essentially an algorithm for recursively splitting the dataset into two subsets which form the subtrees. The purpose is to maximise I[X;Y], where I is the information that the subtree provides, Y is the response variable and X_m, $m = 1, \ldots, M$, are the predictor variables used to construct the purest possible subtrees [10]. Only univariate splits are considered which means that each split depends on the value of only one predictor value. A tree is grown according the following algorithm [11]:

1. Find each predictor's best split: For each predictor, sort its values from the smallest to the largest. Go through each value of the predictor and determine the best splitting point as the one that maximises the splitting criterion (to be defined shortly) if the node is split according to it.
2. Find the node's best split. Among the best splits found in step 1, choose the one that maximizes the splitting criterion.
3. Split the node using its best split found in step 2 if the stopping rules are not satisfied. If a split takes place then repeat the previous steps for both of the resulting new nodes.

At node t, the best split s is chosen by maximising the splitting criterion:

$$\Delta i(s,t) = i(t) - p_L i(t_L) - p_R i(t_R) \tag{1}$$

where:

$$i(t) = \frac{\Sigma n \in \hbar(t) fn(yn - y(t))^2}{\Sigma n \in \hbar(t) fn} \tag{2}$$

with n as an index for the summation of Y, $\hbar(t)$ as the set of observations that fall in node t, f_n as the frequency weight associated with case n, where

$$p_L = N_W(t_L)/N_W)(t), \tag{3}$$

$$p_R = N_W(t_R)/N_W(t) \tag{4}$$

with $N_w(t)$ as the weighed number of cases in node t, $N_w(t_L)$ the number of cases which go to the left after splitting and $N_w(t_R)$ the number of cases which go the right after splitting.

$$N_W(t) = \Sigma_{n \in \hbar(t)} \cdot fn \tag{5}$$

and

$$\dot{y}(t) = \frac{\Sigma n \in \hbar(t) fn \cdot yn}{N_W(t)} \tag{6}$$

[11]. With this splitting criterion a tree with a high number of subtrees can be constructed, which may result in too many terminal nodes. It is possible that there are so many terminal nodes that the final tree overfits the data (rather than providing a representation of the population). To decrease the number of terminal nodes one can prune the tree. There are several criteria for impurity and the tree is grown until one of the criteria for impurity is met. The criterion that is used here is the minimum change in improvement of the tree. The improvement of a split is defined as:

$$\Delta I(S^*, t) = p(t)\Delta i(S^*, t) \tag{7}$$

where $p(t)$ is the probability of a case being in node t. If for the best split s^* of node t the improvement is smaller than the minimum