ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 02/2006

An Economic Model to Evaluate the Contribution of Genetics to Milk Production

Antonio Álvarez y David Roibás

Departamento de Economía

Universidad de Oviedo

Available online at: www.uniovi.es/economia/edp.htm

UNIVERSIDAD DE OVIEDO

DEPARTAMENTO DE ECONOMÍA

PERMANENT SEMINAR ON EFFICIENCY AND PRODUCTIVITY

AN ECONOMIC MODEL TO EVALUATE THE CONTRIBUTION OF GENETICS TO MILK PRODUCTION

Antonio Alvarez♣ and David Roibás♥

Efficiency Series Paper 02/2006

Abstract: In this paper we focus on the role that genetic progress may play in improving milk quality. Despite important genetic advances in dairy, the absence of genetic records in farm management data bases has precluded empirical production models from explicitly accounting for differences in genetics across herds. Our objective is to analyze the contribution of genetics to milk production. Genetics are expected to affect the quantity of milk produced and to influence milk quality by affecting fat and protein content. We first estimate a dairy production function which includes a genetic index of milk quantity as one of the inputs. Next, we analyze the influence of genetics on milk composition, by splitting milk production into protein, fat and other components. The paper explores some modeling issues associated with the specification of the effect of genetics in this multi-output technology framework. In particular, we consider genetic indexes as allocable inputs and the remaining inputs as non-allocable.

Key words: dairy farms, genetic, milk quality, multi-output technology

♣ Corresponding author. Departamento de Economía, Universidad de Oviedo, Spain.
e-mail: alvarez@uniovi.es

♥ Universidad de Oviedo, Spain.

1. Introduction

Major changes in the Common Agricultural Policy (CAP) introduced by the Mid-Term Review pose new challenges to dairy farmers in the European Union. In particular, the reduction in the intervention prices of butterfat and skimmed milk powder is expected to produce a decline in farm level milk prices, which in turn will put downward pressure on the income of dairy farmers[1]. In this environment dairy farmers will need to improve their management capabilities in order to survive.

Given that dairy quotas have been maintained, improving the financial performance of the farms can be based on two strategies: one is to reduce costs; the other, to increase revenues by improving milk quality. Reducing the cost of production depends upon the identification of the sources of inefficiency. The literature on this topic is rather large, with many papers studying the efficiency of dairy farms (e.g. Dawson and Hubbard, 1985; Tauer and Belbase, 1987; Bravo-Ureta and Rieger, 1991). On the other hand, the strategy of improving milk quality has not been the subject of much research. This is somewhat surprising since the dairy industry has adopted multiple-component pricing around the world, giving more value to milk constituents in addition to hygienic quality. Thus, an appropriate strategy for boosting revenues is to increase the levels of milk components. This is the situation of many Spanish dairy farms, where average fat and protein content are around 3.1 % (protein) and 3.7% (fat), lower than in most western European countries. During the period analyzed in the paper (1999-2002) each 0.1% increase in protein (fat) content originates a premium of 0.0042 € (0.003 €) per liter. This means that if farmers could increase protein and fat content by 0.1%, their income would increase by 2.5%., given that the average price of milk during the period was 0.2868 € per liter.

The challenge is how to increase the production of milk components, a topic which has been the subject to much research world-wide. We can not summarize the abundant literature here, but the bottom line is that milk composition can be altered mainly by dietary manipulation or through genetic improvement. While the effect of feeding on milk composition has been widely studied, the effect of genetics has received less attention.

---

[1] Even though the reform introduced a compensation package, the so called 'dairy premium', it is unclear if it will be large enough to compensate the effects of the expected drop in farmgate milk prices.

In this paper we focus on the role that genetic progress may play in improving farmers' incomes. It is widely accepted that tremendous genetic improvements have occurred in the dairy cattle population in recent decades. This progress has been mainly due to the availability of top category sires, the widespread adoption of artificial insemination, and the discovery of new technologies such as embryo transfer. Despite the important evolution in genetics, the absence of genetic records in farm management data bases has precluded empirical models of dairy production from explicitly accounting for differences in genetics across herds.[2]

We use a novel data set of Spanish dairy farms where we have combined genetic records with management data. The sample reports data on 83 dairy farms during the period 1999-2002. The genetic information consists of genetic indexes for the dairy cows from each herd. These genetic indexes are available for a large number of cow characteristics (usually referred to as 'traits'). While attention is often given to non-production traits, we focus on traits related to production. In particular, we will use three cow-specific genetic indexes that reflect the cows' genetic quality with regard to the production of milk, protein and fat.

Our objective is to use this sample in order to analyze the contribution of genetics to milk production and milk components. Genetics are expected to affect milk production in two ways. First, the genetic makeup of a cow will affect her ability to convert feed into a greater quantity of milk. Second, genetics influence milk quality by affecting fat and protein content, thereby generating variation in the value of raw milk.

We first focus on the contribution of genetics to the quantity of milk produced. Since most papers have estimated dairy production functions using a single measure of output such as liters of milk, our first model will be a production function which includes the genetic index for milk quantity as one of the inputs used by the farms. The inclusion of genetics in a production function will allow a more precise evaluation of the technical

---

[2] A few papers have dealt with this issue by using proxies to account for the effect of genetics. For example, Richards and Jeffrey (1996) used the annual expenditure on breeding and veterinary care, Alvarez and Gonzalez (1999) used the price of the most expensive semen dose employed in artificial insemination, while Mbaga et al. (2003) used the weight of the cows.

characteristics of milk production. In particular, we expect to identify how much technical progress is due to factors other than genetic improvement.

Next, we turn to the influence of genetics on milk composition. To analyze this issue we split milk production into three components: protein, fat and other components[3]. We study the effect of genetics in the framework of an output-oriented distance function where inputs are used to produce milk components. The paper explores some modeling issues associated with the specification of the effect of genetics in this multi-output technology framework. In particular, we consider genetic indexes as allocable inputs and the remaining inputs as non-allocable. Allocability implies that producers can decompose the total amount of one input into various parts, and link these parts to each output separately (Lau, 1972; Shumway, Pope and Nash, 1984). That is, the genetic quality of the herd in producing, say, protein is considered to influence only the quantity of protein and not other components.

Our paper makes two major contributions. First, to the best of our knowledge, this is the first paper to explore the effect of genetics on milk production using farm management data[4]. Second, we model milk production as a multi-output technology, where outputs are milk components. Some papers have also considered several outputs, such as milk production and livestock sales (e.g. Tauer and Belbase, 1987), but again to the best of our knowledge we are not aware of published papers that have considered milk components as outputs.[5]

The structure of the paper is the following. Section 2 develops the theoretical model where the production of milk components is specified as a function of both allocable and non-allocable inputs. In section 3 we describe the data set. Section 4 contains the specification of the empirical model. The estimation and results are in Section 5. The implication of the results for management and policy purposed are discussed in section 6. Finally, Section 7 concludes.

[3] The three components are measured in kilograms. ‘Other components’ is the weight of milk after subtracting the weight of fat and protein. Therefore, it includes the weight of water.

[4] One exception is the paper by Weersink et al. (1991) that studied the relationship on dairy farms between profits and breeding decisions, measured by a weighted average of the genetic merit of the sires used in the artificial insemination process on the farms. However, they did not use the genetic level of the cows.

[5] A recent exception is the paper by Cho and Tauer (2006).

2. Model specification

In this section we develop a primal multi-output model of dairy production that takes into account differences in the genetic levels of the herds. Since we have data on several genetic traits we model its influence by including genetic indexes as inputs in the production process. Milk is split into three components (each one associated with a genetic characteristic) and we consider milk production as the joint production of its components. Thus, we model milk production as a multi-output technology where some inputs are used to produce several milk components. This multi-output technology can be represented by means of a transformation function:

$$T(y, x, z) = 0 \qquad (1)$$

where y is the vector of milk components, x is the vector of non-genetic inputs and z is the vector of genetic inputs.

Non-genetic inputs such as labor, cows, feed, etc. can be considered as non-allocable, because they jointly affect the production of every milk component, i.e. it is impossible to split the labor endowment into several parts linking each part to the production of one milk component. On the other hand, we consider genetic inputs as allocable because each genetic characteristic influences only one milk component. We expect that, ceteris paribus, the only effect of an improvement in the genetic quality of the herd in terms of protein production is an increase in protein, with no effect on the production of fat or other components. Thus, the effect of one genetic characteristic on its associated milk component is independent of other genetic characteristics affecting the production of other milk components. These independent effects associated with genetic inputs imply some separability restrictions on the technology that have to be imposed on equation (1) before proceeding to the empirical analysis.

It is worth noting that if for every input vector (x, z) there exists a unique output vector (y) it is possible to describe this technological relationship, imposing the separability assumptions, by means of a set of production functions $y_m = y_m(x, z_m)$. However if the technology shows some possibilities of substitution among outputs then each input vector generates a transformation curve and a set of production functions can not be used to describe this technological relationship.

Following Peterson, Boisvert and Gorter (2002) we impose the separability restrictions in (1) by describing the technology using a combination of a transformation function and a set of production functions, one for each output. The transformation function takes account of the effect of non-allocable inputs on production and of the substitution relationships among different outputs. The production functions are used to capture the effect of allocable inputs on their corresponding outputs.

To carry out the empirical analysis it is necessary to create an equation integrating the set of production functions into the transformation function. To perform this integration we first define a restricted transformation function which exhaustively describes the technology subject to the restriction that the allocable input vector takes some fixed reference value z^0. The restricted transformation function defines a set of transformation curves among outputs, one for each non-allocable input endowment, subject to the restriction that the allocable input vector takes value z^0. Then, we proceed to define a set of incremental production functions which take account of the variations in production that can be generated if the farm applies an allocable input endowment different from z^0. While the restricted transformation function generates a map of transformation curves depending on the non allocable input vector (subject to the restriction $z = z^0$) the set of incremental production functions move each point on this map generating a new one for each endowment of z different from z^0.[6]

Defining y^0 as the output vector attainable using the allocable input endowment z^0 the restricted transformation function can be represented as follows:

$$T^0\left(y^0, x\right) \equiv T\left(y, x, z^0\right) = 0 \tag{2}$$

Figure 1 shows the transformation curve that the technology generates for a firm that produces two outputs using z^0 along with a non-allocable input vector x^1. This function characterizes substitution possibilities among different outputs.

INSERT FIGURE1

[6] It could be necessary to consider some restrictions when choosing z^0. For example, if z_m were essential to produce y_m then choosing z^0 as a null vector would imply that the firm produces zero units of each output and no transformation curves among outputs are generated in the restricted technology. Thus, to allow that the restricted transformation function generates a set of transformation curves it would be necessary to choose z^0 as a strictly positive vector.

Once the restricted technology has been characterized, we proceed to analyze the effect of applying an allocable input vector different from z^0 in the production of milk components. As, ceteris paribus, the variations of z_m with respect to z_m^0 only affect y_m we define a set of incremental production functions, one for each output, to capture these effects. Defining Δz as the difference between the current genetic endowment z and the vector of reference z^0, the set of incremental production functions are:

$$y_m = f_i\left(y_m^0, \Delta z_m, x\right); \quad where\ \Delta z_m = z_m - z_m^0 \qquad (3)$$

Equation (3) establishes the effect of z_i on y_i by assessing how the jump from z_m^0 to z_m originates a variation in production from y_m^0 to y_m. The presence of x in $f_m(\cdot)$ allows the effect of z_m to be non-separable from the non-genetic input endowment. The absence of z_n in (3), on the other hand, assures the separability assumptions associated with allocable inputs.

Figure 2 shows the way in which the functions $f_m(\cdot)$ captures the effect of using an allocable input vector z^1 different to the vector of reference. Once an output vector is selected on the transformation curve in Figure 1, each function in (3) indicates the expansion in its corresponding output originated by Δz_m. This way, for every output vector on the transformation curve exists a new output vector generated by the set of production functions. This new set of output vectors characterizes the transformation curve generated by the technology if the farm were using the allocable input vector z^1 along with the non allocable input vector x^1.

INSERT FIGURE 2

It is important to note that since y^0 is unobserved it is impossible to empirically estimate equations (2) and (3). However, it is possible to integrate the latter into the restricted transformation function to define an equation which describes the whole technology and thereby enables empirical analysis. Solving (3) for y_m^0 we obtain this other set:

$$y_m^0 = f_m^{-1}\left(y_m, \Delta z_m, x\right) \qquad (4)$$

Substituting (4) in (2):

$$T^0\left(f_1^{-1}(y_1, \Delta z_1, x), \ldots, f_M^{-1}(y_M, \Delta z_M, x), x\right) = 0 \qquad (5)$$

Equation (5) represents the unrestricted technology by incorporating the set of inverted production functions in (4) into the transformation function in (2). Additionally, as variables in (5) are observable the equation can be empirically estimated. The separability assumptions associated with genetic inputs are imposed through the set of functions $f_m^{-1}(\cdot)$. Figure 3 depicts the transformation curve generated by moving each point of the restricted transformation curve $T^0(\cdot)$ in Figure 1 according to the effect of applying an allocable input vector z^1 different from the endowment of reference z^0.

INSERT FIGURE 3

Finally, to proceed to empirical analysis we assume that $T^0(\cdot)$ in equation (5) can be expressed by means of an output-oriented distance function. Thus, assuming a translog functional form and imposing that the distance function is homogeneous of degree one in outputs (Kumbhakar and Lovell, 2000) the technology can be expressed by the following equation:

$$\ln y_m^0 = TL\left[\ln\left(\frac{y_n^0}{y_m^0}\right), \ln x\right] \qquad (6)$$

where TL stands for translog functional form. This function will be the basis to generate our empirical analysis. First, we describe the data set.

3. Data

A key feature of this paper is that we combine two data sets. The first one is a set of farm management records which contains data on milk production, milk composition and the input endowments used by the farms. The second data set comes from the Spanish Dairy Herd Improvement Association (CONAFE) and contains genetic indexes at the cow level. Our sample consists of 83 Spanish dairy farms specialized in milk production during the period 1999-2002. The farms represented in the sample are located in Northern Spain and all of them use exclusively Frisian cows. Since some farms were not observed during the whole period, the sample is an unbalanced panel data set of 315 observations. Before analyzing the main characteristics of the farms in the sample we briefly describe the calculation of the genetic indexes.

A genetic index (GI) is an estimate of genetic merit which can be calculated for a wide number of characteristics of dairy cows (usually called traits) including milk, protein, fat, and type traits[7]. These indexes are calculated for bulls and cows using a procedure called the Animal Model[8]. Each GI is given in the units used to measure the trait. For example, the GI for milk is reported in kilograms. GIs are expressed as differences from the breed base, which is equivalent to the genetic merit of the average animal in the population. In fact, the base is defined by setting the average predicted transmitting ability to zero for a group of animals.[9]

The interpretation of the GIs is the following. A GI in a particular trait is the genetic value that an animal transmits to its offspring. The expected breeding value of a daughter is the average of the GIs of her sire and dam. As an example, a cow with a GI of 1500 for milk is expected to produce daughters averaging 250 more kilograms per lactation than daughters of a cow with a GI of 1000.

As stated in the Introduction, we use genetic indexes for three main productive traits: milk, fat, and protein yield. Since the GIs are calculated as differences between the evaluated animal and the reference cow, they can take on negative values (while our empirical model requires taking logs). Thus, we have added to each GI the value of the reference cow in order to make them positive. The genetic indexes of the herd are then calculated as the arithmetic mean of the cows' genetic indexes. Table 1 shows that the genetic quality of the herds in our sample has increased over the sample period. From 1999 to 2002 the index of milk production increased 2.3%, the index of fat production increased by 1.6% and the index of protein production went up by 2.4%. From the

[7] Our genetic indexes are equivalent to the ones used in other countries, such as the PTA (Predicted Transmitting Ability) indexes in the US or the ETAs (Estimated Transmitting Ability) in Canada.

[8] The Animal Model is a system for genetic evaluations that estimates breeding values of bulls and cows at the same time. This system uses production data on all known relatives in calculating a genetic evaluation. As a result, the Animal Model provides more accurate genetic evaluations.

[9] The breed base is occasionally recalculated every 5 years. Because genetic merit of the population improves over time, a recalculation typically causes all GIs to be reduced. This recalculation has no effect on the ranking of animals.

evolution of the indexes it seems that dairy farmers are trying to increase milk and protein production, paying less importance to fat.[10]

INSERT TABLE 1

In our empirical analysis dairy farms are assumed to produce milk (y) using five inputs: labor (x_1), cows (x_2), concentrates (x_3), forage (forage purchases and expenditure on seed, fertilizers, machinery, fuel and land) (x_4) and animal expenses (livestock supplies, breeding and veterinary expenses) (x_5). As for output, we decomposed milk into three main components: kilograms of protein (y_1), kilograms of fat (y_2) and kilograms of other components (y_3). Finally, the genetic indexes used in empirical analysis are the genetic index of kilograms of protein (z_1), the genetic index of kilograms of fat (z_2) and the genetic index of kilograms of milk (z_3). Table 2 reports some descriptive statistics of the distribution of outputs and inputs of the dairy farms in the sample during the period analyzed.

INSERT TABLE 2

It must be noted that while the scale of the farms varies substantially (e.g., cows ranges from 11 to 88) the genetic quality of the herds are more similar across farms. The reason for this limited variation in genetics could be that the farmers in the sample are voluntarily associated to CONAFE, implying that all of them are trying to achieve a herd with high genetic merit.

4. Empirical model

We estimate several models. In the first model our objective is to assess the importance of genetics on the quantity of milk produced. For this purpose we estimate a translog production function where the genetic index of milk production (z_3) enters as just another input.

[10] One reason for paying less importance to fat is that milk quotas are given in kilograms of fat. Therefore, boosting fat can lead to exceeding the quota, which carries severe penalties.

$$\ln y = \alpha_i + \sum_{j=1}^{5} \beta_j \ln x_j + \frac{1}{2}\sum_{j=1}^{5}\sum_{k=1}^{5} \beta_{jk} \ln x_j \ln x_k + \gamma_3 \ln z_3 + \gamma_{33}(\ln z_3)^2 + \sum_{j=1}^{5} \eta_{3j} \ln z_3 \ln x_j + \alpha_{2000}T_{2000} + \alpha_{2001}T_{2001} + \alpha_{2002}T_{2002} \quad (7)$$

where α_i, β_j, β_{jj}, γ_j and η_{jj} are parameters to be estimated and the T_t are dummy variables which equal one in the year indicated by the subscript and zero otherwise

In the second model we analyze the influence of genetics in milk composition. We consider milk production as the multi-output production of its components, namely protein, fat and other components. The analysis is carried out estimating a model which combines an output-oriented translog distance function with a set of production functions, one for each milk component. The output-oriented distance function relates the output attainable using the genetic input endowment of reference (z^0) and the non allocable inputs:

$$\ln y_1^0 = \alpha_i + \sum_{j=1}^{5} \beta_j \ln x_j + \frac{1}{2}\sum_{j=1}^{5}\sum_{k=1}^{5} \beta_{jk} \ln x_j \ln x_k + \sum_{m=2}^{3} \theta_m \left(\ln y_m^0 - \ln y_1^0\right) + \frac{1}{2}\sum_{m=2}^{3}\sum_{n=2}^{3} \theta_{mn} \left(\ln y_m^0 - \ln y_1^0\right)\left(\ln y_n^0 - \ln y_1^0\right) + \sum_{m=2}^{3}\sum_{j=1}^{5} \mu_{mj} \left(\ln y_m^0 - \ln y_1^0\right)\ln x_j + \alpha_{2000}T_{2000} + \alpha_{2001}T_{2001} + \alpha_{2002}T_{2002} \quad (8)$$

where α_i, β_j, β_{jj}, θ_j, θ_{jj}, and μ_{jj} are the parameters to be estimated.

The production functions in (4) are also modeled using a translog functional form generating a set of three production functions as:

$$\ln y_m = \ln y_m^0 + D_m \left(\gamma_m \ln \Delta z_m + \frac{1}{2}\gamma_{mm} (\ln \Delta z_m)^2 + \sum_{j=1}^{5} \eta_{mj} \ln \Delta z_i \ln x_j \right) \quad (9)$$

where γ_m, γ_{mm} and η_{mj} are the parameters to be estimated and the D_i are dummy variables defined as:

$$\begin{aligned} D_m &= 1 \quad if\ z_m > z_m^0 \\ D_m &= 0 \quad if\ z_m = z_m^0 \end{aligned} \quad (10)$$

It was mentioned above that allocable inputs could be essential in production, in which case choosing z^0 as a null vector would cause some identification problems. We therefore define each element of z^0 using the sample minimum of each genetic index.

From (9) and (10):

$$\ln y_m^0 = \ln y_m - D_m\left(\gamma_m \ln \Delta z_m + \frac{1}{2}\gamma_{mm}(\ln \Delta z_m)^2 + \sum_{j=1}^{5}\eta_{mj} \ln \Delta z_m \ln x_j\right) \qquad (11)$$

Substituting the set of equations referred in expression (11) into the equation (8) and solving for ln y_1 we obtain equation (12) that, adding the corresponding error term, can be econometrically estimated.

$$\begin{aligned}
\ln y_1 = {} & \alpha_i + \sum_{j=1}^{5}\beta_j \ln x_j + \frac{1}{2}\sum_{j=1}^{5}\sum_{k=1}^{5}\beta_{jk} \ln x_j \ln x_k + \\
& \sum_{m=2}^{3}\theta_m\left(\begin{array}{l}\left[\ln y_m - D_m\left(\gamma_m \ln \Delta z_m + \frac{1}{2}\gamma_{mm}(\ln \Delta z_m)^2 + \sum_{j=1}^{5}\eta_{mj} \ln \Delta z_m \ln x_j\right)\right] - \\ \left[\ln y_1 - D_1\left(\gamma_1 \ln \Delta z_1 + \frac{1}{2}\gamma_{11}(\ln \Delta z_i)^2 + \sum_{j=1}^{5}\eta_{1j} \ln \Delta z_i \ln x_j\right)\right]\end{array}\right) + \\
& \frac{1}{2}\sum_{m=2}^{3}\sum_{n=2}^{3}\theta_{mn}\left(\begin{array}{l}\left[\ln y_m - D_m\left(\gamma_m \ln \Delta z_m + \frac{1}{2}\gamma_{mm}(\ln \Delta z_m)^2 + \sum_{j=1}^{5}\eta_{mj} \ln \Delta z_m \ln x_j\right)\right] - \\ \left[\ln y_1 - D_1\left(\gamma_1 \ln \Delta z_1 + \frac{1}{2}\gamma_{11}(\ln \Delta z_1)^2 + \sum_{j=1}^{5}\eta_{1j} \ln \Delta z_1 \ln x_j\right)\right]\end{array}\right) \\
& \left(\begin{array}{l}\left[\ln y_n - D_n\left(\gamma_n \ln \Delta z_n + \frac{1}{2}\gamma_{nn}(\ln \Delta z_n)^2 + \sum_{j=1}^{5}\eta_{mj} \ln \Delta z_n \ln x_j\right)\right] - \\ \left[\ln y_1 - D_1\left(\gamma_1 \ln \Delta z_1 + \frac{1}{2}\gamma_{11}(\ln \Delta z_1)^2 + \sum_{j=1}^{5}\eta_{1j} \ln \Delta z_1 \ln x_j\right)\right]\end{array}\right) + \\
& \sum_{m=2}^{3}\sum_{j=1}^{5}\mu_{mj} \ln x_j\left(\begin{array}{l}\left[\ln y_m - D_m\left(\gamma_m \ln \Delta z_m + \frac{1}{2}\gamma_{mm}(\ln \Delta z_m)^2 + \sum_{j=1}^{5}\eta_{mj} \ln \Delta z_m \ln x_j\right)\right] - \\ \left[\ln y_1 - D_1\left(\gamma_1 \ln \Delta z_1 + \frac{1}{2}\gamma_{11}(\ln \Delta z_1)^2 + \sum_{j=1}^{5}\eta_{1j} \ln \Delta z_1 \ln x_j\right)\right]\end{array}\right) \qquad (12) \\
& + D_1\left(\gamma_1 \ln \Delta z_1 + \frac{1}{2}\gamma_{11}(\ln \Delta z_1)^2 + \sum_{j=1}^{5}\eta_{1j} \ln \Delta z_1 \ln x_j\right) + \\
& \alpha_{2000}T_{2000} + \alpha_{2001}T_{2001} + \alpha_{2002}T_{2002}
\end{aligned}$$

5. Estimation and results

We estimated equation (7) using the fixed-effects estimator.[11] The econometric program was TSP. The explanatory variables were divided by the sample geometric mean, so the first order parameters can be interpreted as elasticities at the sample geometric means. The results with and without the genetic index for milk production are displayed in Table 3.

INSERT TABLE 3

As can be seen, both estimations are quite similar. The elasticity of scale equals 0.878 when the genetic index is included and 0.873 without the index. In both cases the Wald test rejects constant returns to scale at the 10% significance level but not at the 5% level. The main difference between the two estimations is in the first order coefficient of labor (β_1). When the genetic index is included in the estimation the elasticity of labor is not significantly different from zero,[12] while it is negative and significant at the 10% level if the genetic index is left out. Thus, the omission of genetics in the empirical analysis suggests a misspecification bias that can lead to misleading conclusions regarding the productivity of some inputs.

Analyzing the effect of genetics on milk production, three of the seven parameters interacting with the genetic index are significantly different from zero. An F test on the set of parameters that interact with the genetic index ($F_{(7,202)}$= 2.443) allows us to reject the null hypothesis that the genetic index does not influence milk production. However, the elasticity of the genetic index evaluated at the sample geometric mean, while positive, is not significantly different from zero.

Regressing the logarithm of the genetic index of milk production on fixed and time effects yields an R^2 of 0.36. Thus, the non-significance of the genetic index may be due to its correlation with the fixed and time effects. It turns out that if we exclude the time dummies from the estimation, the elasticity of the genetic index is clearly positive and significant.

---

[11] We use the fixed-effects estimator because it is consistent even if the individual effects are correlated with the explanatory variables. On the other hand the fixed-effects estimator imposes that inefficiency is time invariant. However, given the length of the panel this restriction can be acceptable.

[12] The non-significance of labor is not unusual in studies which use family farms with very little hired labor since the sampling variability is very small (see, for example, Ahmad and Bravo-Ureta).

The time effects are positive and significant in both models. From 1999 to 2002 the calculated productivity growth is about 6.9% when genetics is included and 7.8% when it is excluded[13]. This result suggests that there are other important reasons that explain technical progress aside from genetic evolution.

The estimated fixed effects can be interpreted as indicators of persistent technical efficiency.[14] Table 4 shows some descriptive statistics of the technical efficiency index estimated with and without genetics. Mean efficiency increases from 71% to 75% when genetics is included in the analysis. However, on close inspection of the data it is observed that the differences in genetics across farms are very small in each period. The difference between the maximum and the minimum genetic endowment year to year ranges from 12.5% in 1999 to 7.5% in 2002. Thus, although differences in genetics generate differences in observed productivity, it is clear that they are not able to explain the whole difference. This suggests that most of the differences in farm productivity correspond to pure technical inefficiency rather than to genetics.

INSERT TABLE 4

Our second empirical analysis deals with the influence of genetics on milk composition. Equation (12) has been estimated using the Within estimator. Tables 5, 6, 7 and 8 show the estimation of the parameters in equation (12). The parameters which describe the restricted technology are reported in Table 5. The parameters which characterize the production functions are reported in tables 6, 7 and 8.

INSERT TABLE 5

The first order coefficients of non-genetic inputs (β_1 to β_5) are similar to those in the production function including the genetic index of milk quantity. The output elasticities with respect to cows, concentrates and forage expenses are positive and significant and take values close to those obtained in the production function, while the elasticity of labor is again not significantly different from zero. The elasticity of animal expenses,

[13] The interpretation of the coefficient of a dummy variable, say γ, when the dependent variable is in logs is ($e^{\gamma} - 1$) (see Suits 1983).

[14] Schmidt and Sickles (1984) propose measuring technical efficiency as: $TE_i = \exp(\alpha_i - \max\alpha_j)$.

however, is positive but not significantly different from zero. The scale elasticity of 0.94 is also similar to the one obtained in the previous regression and according to a Wald test is not significantly different from one so constant returns to scale can not be rejected. The coefficients associated with the time dummy variables also have estimated values similar to those obtained in the previous regression (the accumulated technical change is 6.5%, which is close to the 6.9% found above). The first order parameters associated with fat and other components are negative and significant, showing the existence of substitution between fat and other components with protein in the restricted technology. Finally, the efficiency levels obtained in this multi-output analysis are similar to those obtained with the production function. The mean of the estimated efficiency is 74% while it was 75% in the former estimation.

As can be seen from the above results, the technological characteristics related to non-genetic inputs are quite similar regardless of whether milk is decomposed into several components or not. This is an important result because it confirms that splitting milk into various components to characterize the technology is a method that does not alter the nature of the results obtained. However, the decomposition of milk into protein, fat and other components allows the influence of genetic factors on milk composition to be studied. Tables 6, 7 and 8 show the estimation of the parameters contained in expressions (11).

INSERT TABLE 6

As we can see, genetics clearly influences the protein content of milk. Six out of seven parameters interacting with the genetic index of protein production are significant, indicating that the genetic quality of the herd is important in order to produce protein. Moreover, according to our results this effectiveness increases with labor and the number of cows but decreases with concentrates and forage expenses.

INSERT TABLE 7

Table 7 shows that five out of seven parameters interacting with the genetic index of fat production are significant, indicating that the genetic quality of the herd affects fat production. Additionally, the effectiveness of genetic quality improves with labor and number of cows but decreases with concentrates, so concentrate use lowers the productivity differences across herds due to genetics.

INSERT TABLE 8

Only two parameters associated with the index of milk production are significant, and γ_3, the direct impact of this index, is not significant. It should be noted that this index is related to the production of raw milk, so the relationship between the index and the output 'other components' may be weak. Nevertheless, the significant parameters show, once again, that the effect of genetics is higher in large farms than in small ones.

It is clear from our empirical results that genetics contributes to produce both more and more valuable raw milk and that its contribution is far from being Hicks-neutral. In the next section we expand on these results.

6. Discussion

Our results indicate a significant relationship between the productivity of genetics and the size of the farm. The parameter η_{32} in Table 3, corresponding to the production function estimation, shows that the productivity of genetics increases with the size of the herd. In a similar vein, the significance of the parameters η_{11}, η_{12}, η_{21}, n_{22}, η_{31} and η_{32} in Tables 6, 7 and 8, corresponding to the milk components analysis, indicates that the effect of genetics grows with labor and the number of cows.

We believe that this result may be driven by differences in management across farms. Our interpretation is that high-genetic herds are more difficult to manage than low-genetic herds, which implies that in order to exploit the potential of high-genetic herds farmers need to improve their management skills. On the other hand, it has been shown that there is a positive relationship between efficiency and farm size. The argument, as developed by Jovanovic (1982) and others, is that an improvement in the farmer's management skills will result in an increase in the size of the farm. Alvarez and Arias (2004) have confirmed the existence of this relationship in a sample of Spanish dairy farms.

In our model management is assumed to be picked up by the fixed effects. In order to explore whether larger farms exhibit higher efficiency scores than small farms we have regressed the efficiency scores obtained in the production function which included genetic indexes and those obtained in the multi-output analysis on some of the variables that can be used to characterize the size of the farms: the individual means of labor and cows. The results are presented in Table 9.

INSERT TABLE 9

The analysis suggests that the bigger farms, that take more advantage of genetics, are better managed than the small ones. Thus, it seems that managing a herd with high genetic merit is not a simple task and only the better managed farms are able to fully exploit its productive potential. Therefore, agricultural extension may play an important role in order for farmers to be able to exploit the potential of genetic improvement.

To evaluate the return to genetics we calculated the protein, fat, other components and milk production as well as the percentage of protein and fat in milk of a representative farm under three different scenarios. The representative farm is defined as an efficient farm which uses the sample geometric mean of non-genetic inputs. The scenarios are related to the use of Low, Average or High Genetic Herds. We set the genetic potential of the Average Genetic Herd equal to the sample geometric mean in the three genetic indexes. The Low (High) Genetic Herd has genetic indexes equal to the sample geometric mean minus (plus) one standard deviation.

INSERT TABLE 10

Inspection of Table 10 shows the importance of genetics in milk production. The High Genetic Herd representative farm would produce 9820 liters of milk more than a farm which uses a Low Genetic Herd. While this difference is not too large the impact of genetics on milk quality is rather important. The percentage of protein goes from 2.95% for the farm with the Low Genetic Herd to 3.46% if the genetic level was high. On the other hand, fat content increases from 3.4% to 4.7% when the farm passes from using a Low Genetic Herd to use a High Genetic Herd. Thus, according to our simulation Spanish dairy farmers should not expect important increases in milk production from investing in genetics. However, investing in genetics can be very profitable because of its effect on fat and protein content, which allows a higher price to be obtained. It is worth noting that, taking as reference the average price of milk during the period (0.2868 € per liter) and the premium-penalty structure that the Spanish dairy industry establishes depending on protein and fat content,[15] passing from the Average Genetic Herd to the High Genetic Herd generates a 15 % increase in income, where 1.5 % is due to the increase in production and the rest is due to the effect of the milk composition.

7. Concluding remarks

Important genetic progress has taken place in milk production in recent decades. However, the absence of farm management data bases with information about the genetic quality of the herds has precluded empirical analysis of the impact of genetics on milk production and milk components. This paper fills this gap in the literature by providing a first attempt at studying the productivity of genetics using farm management data. We are able to accomplish this objective by using a sample of Spanish dairy farms that contains genetic records in addition to the typical management data.

Genetics influence milk production in two ways: by increasing the amount of milk produced and affecting its composition. For this reason, in the empirical section we estimate two models. First, we include genetics as a normal input in a production function. Second, we split milk into three main components (protein, fat, and 'other components') and develop an empirical model which is able to capture the influence of both genetic (allocable) and non-genetic (non-allocable) inputs.

The empirical results show several interesting characteristics of the effect of genetics on milk production. The effect of genetics on milk composition, which implies more valuable milk, is stronger than the impact on the quantity of milk produced. Another interesting result is that increased use of concentrates lowers the effectiveness of genetics in producing milk richer in protein and fat. Finally, it is worth noting that genetics appears to be more effective in larger than in smaller farms. Since large farms are more efficient than small farms, this result suggests that farms with better

[15] Each 0.1% of protein content above (below) the 3.1% of reference generates a premium (penalty) of 0.0042 € per liter. On the other hand, each 0.1% of fat content above (below) the 3.7% of reference generates a premium (penalty) of 0.003 € per liter.

management are able to exploit the genetic potential of the herd to a much larger extent than farms with lower management ability. Thus, it seems that managing a herd with high genetic potential is more complicated than managing a genetically lower herd. This suggests that agricultural extension may play an important role in assisting farmers in exploiting the herd's genetic potential.

Our results show a potential to shift milk composition by genetic selection. As component pricing becomes more standard in the industry, production of milk components, particularly protein, appears to be a sensible strategy to be implemented by dairy farms.

As a final remark it is important to note that while genetic improvement appears as an investment that increases productivity and revenues, it also poses some additional costs. Thus, the issue is how much to invest in genetic improvement. The question is of great importance, since raising milk production and quality by means of genetic improvement can, under certain market conditions, have a negative impact on net revenues. Even though some papers have dealt with this issue (e.g., Wilcox et al., 1981; Shumway et al., 1987) it is clear that more additional work is needed to study the optimal allocation of genetics and its net contribution to farmers' profits.

References

Ahmad, M., and B. Bravo-Ureta, 1995. “An Econometric Analysis of Dairy Output Growth.” American Journal of Agricultural Economics 77: 914-921.

Alvarez, A. and Arias, C. 2004. “Technical Efficiency and Farm Size: A Conditional Analysis”, Agricultural Economic 30: 241-250.

Alvarez, A. and E.González. 1999. "Using Cross-section Data to Adjust Technical Efficiency Indexes Obtained with Panel Data." American Journal of Agricultural Economics 81: 894-901.

Bravo-Ureta, B. E. and L. Rieger. 1991. “Dairy Farm Efficiency Measurement Using Stochastic Frontiers and Neoclassical Duality.” American Journal of Agricultural Economics 73: 421-28.

Cho, J. and L. Tauer. 2006. “Relationships between Milk Composition and Farm Business Characteristics: SUR Estimation of Production versus an Output Distance Function.” Paper presented at the North American Productivity Workshop New York University.

Dawson, P. and L. Hubbard. 1985. “Management and Size Economies in the England and Wales Dairy Sector.” Journal of Agricultural Economics 38: 27-37.

Kumbhakar, S. and C. A. K. Lovell. 2000. Stochastic Frontier Analysis, Cambridge University Press

Jovanovic, B. 1982. “Selection and the Evolution of Industries.” Econometrica, 50: 649-670.

Lau, L. J. 1972. “Profit Functions of Technologies with Multiple Inputs and Outputs.” Review of Economics and Statistics 54: 281-289.

Mbaga, M.D., R. Romain, , B. Larue, and L. Lebel. 2003. "Assessing Technical Efficiency of Québec Dairy Farms." Canadian Journal of Agricultural Economics 51: 121-137 .

Peterson, J. M., R. N. Boisvert, and H. Gorter. 2002. “Environmental Policies for a Multifunctional Agricultural Sector in Open Economies.” European Review of Agricultural Economics 29: 423-443.

Richards, T.J. and S.R. Jeffrey. 1996. “Cost and Efficiency in Alberta Dairy Production.” Rural Economy Staff Paper 96-13, University of Alberta.

Schmidt, P. and R. Sickles. 1984. "Production Frontiers and Panel Data." Journal of Business and Economic Statistics 2: 367-74.

Shumway, R., R. D. Pope, and E. K. Nash. 1984. “Allocatable Fixed Inputs and Jointness in Agricultural Production: Implications for Economic Modelling.” American Journal of Agricultural Economics 66: 72-78.

Shumway, R., Blake, R., D. Leatham, and M. Tomaszewski. 1987. “Assesing the comparative economics of a biotechnology: artificial insemination dairy sires.” Agribusiness 3: 207-220.

Suits, D. 1983. “Dummy variables: mechanics versus interpretation.” Review of Economics and Statistics 66: 177–180.

Tauer, L. and K. Belbase. 1987. “Technical Efficiency of New York Dairy Farms.” Norhteastern Journal of Agricultural and Resource Economics 16: 10-16.

Weersink, A., W. Howard, J. Dekkers, , and M. Lohuis. 1991. “Sire Selection Decisions and Farm Profitability Relationships for Ontario Dairy Farmers.” Journal of Dairy Science. 74: 3994-40001.

Wilcox, M., R Shumway, R.W. Blake, and M.A. Tomaszewski. 1984. “Selection of artificial insemination sires to maximize profits” Journal of Dairy Science 67: 2407-419.

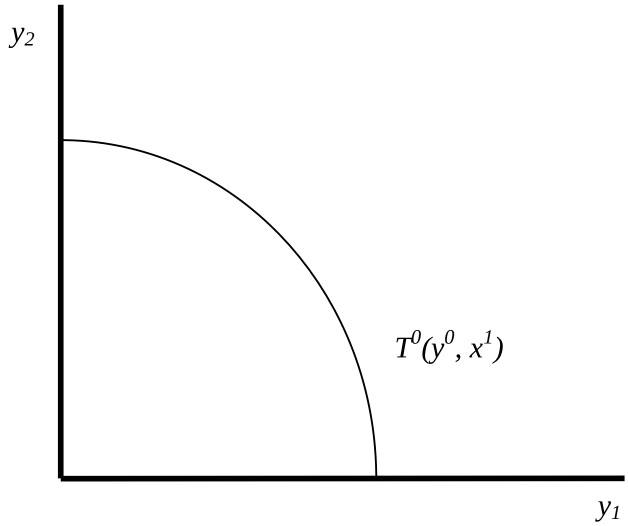

Figure 1. Restricted transformation function

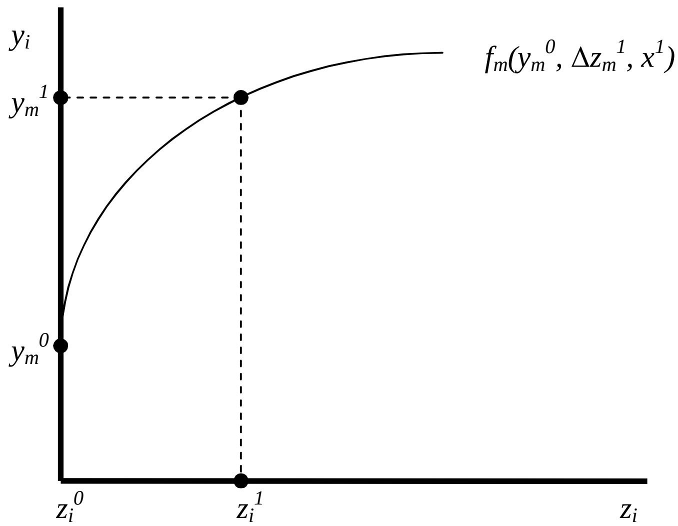

Figure 2.Effect of Δz_m^1 in the production of y_m

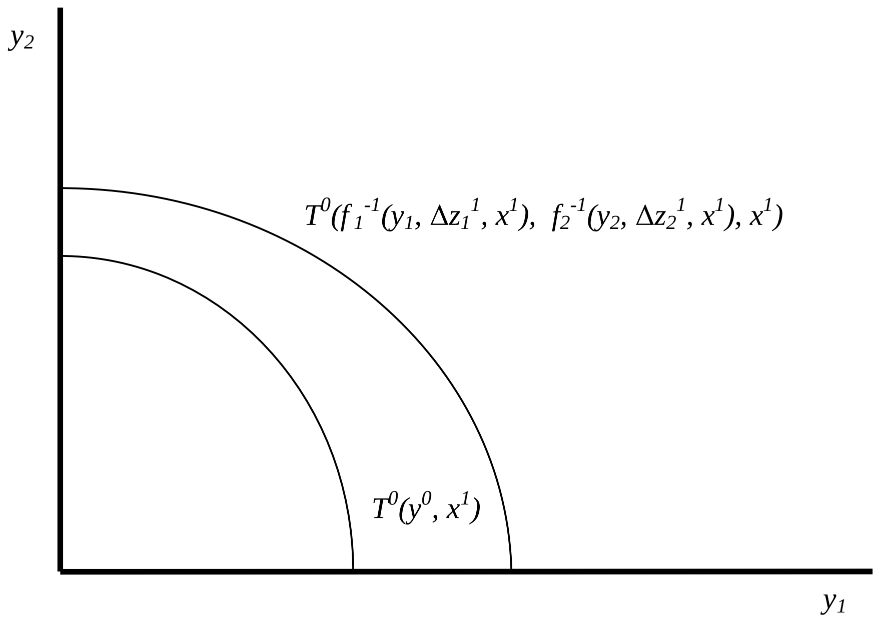

Figure 3. Production technology

Table 1. Evolution of the modified genetic indexes

	Milk GI (kg)	Fat GI (kg)	Protein GI (kg)
1999	8782	318	276
2000	8859	320	279
2001	8918	321	281
2002	8986	323	283

Table 2. Summary Statistics of the Dairy Farms (N=83)

	Mean	St. Dev.	Max.	Min.
Milk Quantity (liters)	287241	147905	773632	64298
Protein (kg)	9102	4895	25649	1946
Fat (kg)	10881	5588	28962	2515
Labor	2.03	0.55	3.50	1
Cows	38.41	15.41	88	11
Concentrates (kg)	141896	80747	417695	23009
Forage Expenses (euros)	18692	12585	76505	1387
Animal Expenses (euros)	3556	2747	15253	171
Genetic Index for Protein Yield	280	4.82	293	263
Genetic Index for Fat Yield	321	5.87	334	304
Genetic Index for Milk Yield	8890	142	9357	8178

Table 3. Estimation of the production function with and without genetics

		With genetic index		Without genetic index	
Variable	Param	Coef.	t-Statistic	Coef.	t-Statistic
Labor	β_1	-0.079	-1.387	-0.105*	-1.865
Cows	β_2	0.623*	9.759	0.658*	10.375
Concentrate	β_3	0.217*	6.015	0.209*	5.709
Forage Expenses	β_4	0.082*	3.587	0.086*	3.735
Animal Expenses	β_5	0.035*	2.397	0.026*	1.763
Labor x Labor	β_{11}	-0.442*	-2.178	-0.421*	-2.060
Cows x Cows	β_{22}	-0.343	-0.921	-0.250	-0.683
Conc. x Conc.	β_{33}	0.073	0.515	0.094	0.657
Forage$ x Forage$	β_{44}	0.073	1.303	0.065	1.143
Animal$ x Animal$	β_{55}	0.025	1.076	0.012	0.526
Labor x Cows	β_{12}	0.415*	1.672	0.263	1.077
Labor x Conc.	β_{13}	0.054	0.366	0.034	0.235
Labor x Forage$	β_{14}	-0.122	-1.582	-0.099	-1.311
Labor x Animal$	β_{15}	-0.025	-0.494	0.016	0.348
Cows x Conc.	β_{23}	-0.024	-0.140	-0.028	-0.161
Cows x Forage$	β_{24}	-0.129	-1.225	-0.105	-1.013
Cows x Animal$	β_{25}	0.025	0.352	0.031	0.450
Conc. x Forage$	β_{34}	0.018	0.272	-0.022	-0.333
Conc. x Animal$	β_{35}	-0.059	-1.279	-0.050	-1.117
Forage$ x Animal$	β_{45}	0.046	1.540	0.010	0.359
T_{2000}	α_{2000}	0.030*	2.433	0.034*	3.070
T_{2001}	α_{2001}	0.040*	2.454	0.047*	3.727
T_{2002}	α_{2002}	0.067*	3.078	0.075*	4.790
Milk genetic index	γ_3	0.232	0.374	----	----
Milk GI x Milk GI	γ_{33}	26.446	0.862	----	----
Milk GI x Labor	η_{31}	1.170	0.660	----	----
Milk GI x Cows	η_{32}	4.386*	2.133	----	----
Milk GI x Conc.	η_{33}	-0.742	-0.545	----	----
Milk GI x Forage$	η_{34}	-2.242*	-2.476	----	----
Milk GI x Animal $	η_{35}	-1.740*	-2.504	----	----
		R^2	99 %	R^2	99 %

* indicates significance at the 10% level.

Table 4. Technical Efficiency distribution

	Mean	Std. Dev.	Max	Min
Efficiency with genetics	0.75	0.090	1	0.53
Efficiency without genetics	0.71	0.093	1	0.49

Table 5. Estimated distance function

Variable	Param	Coef.	t-Stat.	Variable	Param	Coef.	t-Stat.
Fat	θ_2	-0.156*	-1.79	Cows x Conc.	β_{23}	-0.100	-0.52
Other comp.	θ_3	-0.945*	-7.02	Cows x Forage$	β_{24}	0.061	0.55
Labor	β_1	-0.045	-0.68	Cows x Animal$	β_{25}	0.110	1.48
Cows	β_2	0.648*	9.62	Conc. x Forage$	β_{34}	-0.103	-1.38
Conc.	β_3	0.248*	6.25	Conc. x Animal$	β_{35}	-0.012	-0.24
Forage$	β_4	0.068*	2.41	Forage$ x Animal$	β_{45}	-0.008	-0.26
Animal$	β_5	0.017	1.06	Fat x Labor	μ_{21}	1.090*	2.90
Fat x Fat	θ_{22}	1.210	1.54	Fat x Cows	μ_{22}	-0.756*	-2.06
O. comp.x O. comp.	θ_{33}	2.735	1.46	Fat x Conc.	μ_{23}	0.722*	2.54
Fat x O. comp.	θ_{23}	1.537*	2.30	Fat x Animal$	μ_{24}	-0.179	-1.02
Labor x Labor	β_{11}	-0.022	-0.09	Fat x Forage$	μ_{25}	0.036	0.37
Cows x Cows	β_{22}	-0.452	-1.12	O. comp.x Labor	μ_{31}	-0.651	-1.30
Conc. x Conc.	β_{33}	0.125	0.89	O. comp.x Cows	μ_{32}	-1.421*	-2.38
Forage$ x Forage$	β_{44}	0.0439	0.76	O. comp.x Conc.	μ_{33}	1.393*	2.73
Animal$ x Animal$	β_{55}	-0.001	-0.06	O. comp.x Animal$	μ_{34}	-0.182	-0.92
Labor x Cows	β_{12}	-0.102	-0.38	O. comp.x Forage$	μ_{35}	-0.256	-1.41
Labor x Conc.	β_{13}	0.383*	2.43	T_{2000}	α_{2000}	0.036*	2.89
Labor x Forage$	β_{14}	-0.122	-1.49	T_{2001}	α_{2001}	0.042*	2.59
Labor x Animal$	β_{15}	-0.032	-0.59	T_{2002}	α_{2002}	0.063*	2.8
	R^2	99 %					

Table 6. Parameters of the protein production function

Variable	Parameter	Coef.	t-Statistic
Protein genetic index	γ_1	0.331*	4.23
Protein GI x Protein GI	γ_{11}	0.340*	2.39
Protein GI x Labor	η_{11}	0.553*	2.58
Protein GI x Cows	η_{12}	0.773*	2.92
Protein GI x Concentrate	η_{13}	-0.335*	-1.85
Protein GI x Forage$	η_{14}	-0.289*	-2.13
Protein GI x Animal$	η_{15}	-0.018	-0.23

Table 7. Parameters of the fat production function

Variable	Parameter	Coef.	t-Statistic
Fat genetic index	γ_2	0.497*	3.07
Fat GI x Fat GI	γ_{22}	0.431*	2.51
Fat GI x Labor	η_{21}	0.997*	2.97
Fat GI x Cows	η_{22}	0.489*	1.77
Fat GI x Concentrate	η_{23}	-0.681*	-2.45
Fat GI x Forage$	η_{24}	-0.179	-1.19
Fat GI x Animal$	η_{25}	0.215	1.60

Table 8. Parameters of the other components allocable production function

Variable	Parameter	Coef.	t-Statistic
Milk genetic index	γ_3	0.025	0.33
Milk GI x Milk GI	γ_{33}	-0.292	-0.92
Milk GI x Labor	η_{31}	0.517*	1.97
Milk GI x Cows	η_{32}	0.532*	1.69
Milk GI x Concentrate	η_{33}	-0.288	-1.35
Milk GI x Forage$	η_{34}	0.002	0.01
Milk GI x Animal$	η_{35}	-0.034	-0.32

Table 9. Technical efficiency analysis

	Efficiency estimated in the production function		Efficiency estimated in the distance function	
	Estimated Coefficient	t-statistic	Estimated Coefficient	t-statistic
Constant	0.534	18.57	0.569	18.37
Individual mean of labor	0.072	3.79	0.073	3.60
Individual mean of cows	0.0019	2.88	0.0013	1.83
R^2	0.47		0.38	

Table 10. Simulated farms for three different genetic levels

	Low Genetic Herd	Average Genetic Herd	High Genetic Herd
Milk	317191	322467	327011
Protein	9363	10126	11331
Fat	10631	12228	15228
Other Comp.	297195	300112	300451
% Protein	2.95	3.14	3.46
% Fat	3.35	3.79	4.65