

A study on the influencing factors of dairy farming returns in China based on regression analysis

Yilin Ma*

College of Economics and Management, China Agricultural University, Beijing, China, 100083

* Corresponding author Email: mylin0107@163.com

Abstract. Since the reform and opening up, China's dairy industry has continued to improve, as reflected in the significant increase in the scale, standardization, and modernization of dairy farming. At the same time, dairy companies are experiencing a steady acceleration of production changes and technological innovation, and the consumption of dairy products is gradually growing, making milk popular as a necessary food for daily life. As the starting point of the whole dairy industry chain, dairy farming is the fundamental link in the development of this industry. Therefore, analyzing the profitability of dairy farming in China and studying the factors influencing the profitability of farming is essential to ensure the sustainable, efficient, and high-quality development of China's dairy industry. Based on the data released by the National Compendium of Costs and Benefits of Agricultural Products, this study uses regression analysis with the help of the theory of costs and benefits to study the influencing factors of the benefits and puts forward constructive suggestions for different modes of dairy farming to reduce costs and increase benefits.

Keywords: Dairy Farming, Cost Benefit, Descriptive Statistics, Multivariate Regression Analysis.

1. Introduction

The dairy industry is an iconic and strategic industry related to national health, and the healthy and sustainable development of dairy farming is the basic condition for the healthy development of the whole dairy industry. The state and local governments at all levels attach great importance to the transformation of dairy production and operation mode, and the construction of the milk source base is gradually developing in the direction of intensification, standardization, and scale [1]. Dairy production has gradually changed from rough growth to quality and efficiency, especially since the trend of dairy farming scale gradually appeared [2]. However, up to now, China's per capita milk consumption is about 38.3 kilograms, which is only equivalent to 1/3 of the nutritional recommendations, and less than 1/2 of the Asian average [3]. According to statistics, the investment in dairy farming accounts for more than 60% of the industrial chain, but the profit is less than 20%. From the perspective of the entire industrial chain, dairy farming, dairy processing, and dairy distribution three link an input ratio of about 7.5:1.5:1, while the profit ratio of about 1:3.5:5.5 [4], dairy farming links the highest cost, the lowest income, input, and output is seriously inverted. Especially in the breeding link, the profit of breeding cows is low compared with other links in the industry chain, and there are obvious problems such as the unit cost of dairy farming increasing year by year, the ups and downs of raw material milk prices leading to unstable cost margins of farms, and the benefits of large-scale farms with large-scale mass production are not obvious.

According to foreign studies, the development of dairy farming in developed countries has got rid of relying solely on expanding the scale of farming to increase the efficiency of dairy farming, but the use of their own high-end, leading science and technology to greatly promote the rapid development of dairy farming science and dairy farming income increase. Ogunyinka and Ajibefun applied the Tobit model to analyze and found that the main body of the educational level of dairy farmers, the investment in land, and participation in dairy training and publicity are important factors affecting the technical efficiency of dairy farming [5]. investment and participation in dairy-related training and publicity are important factors that affect the technical efficiency of dairy farming [5]. Domestic research on dairy farming is still at the stage of exploring various aspects. Based on the cost-benefit theory, Honghua also put forward rationalization suggestions for optimizing the cost-

control system of large-scale farms [6], and Xing Zhang analyzed the dairy farming industry from the aspects of feeding, epidemiological diseases, industrial scaling, and industrial structure renewal, and finally came up with the improvement strategy of dairy farming [7].

To summarize, the current research on dairy farming is mainly qualitative analysis, and there are fewer studies on different types of dairy farming. Therefore, this study investigates the income of dairy farming under different types of farming in China and analyzes quantitatively the influence of different influencing factors on the net profit of dairy farming by studying the influence of two major types of factors, namely, total cost item and total output value item, and proposes specific methods to improve the income of farming according to the results of the study, so as to promote the realization of high level of "cost reduction and income enhancement" of dairy farming in China as a whole. This is to promote China's dairy farming to realize a high level of "cost reduction and income increase", and to promote the sustainable, healthy, and efficient development of dairy farming of different scales.

2. Regression analysis of the factors influencing the profitability of dairy farming in China

2.1. Multivariate correlation analysis

Based on the existing literature and related experience, it is believed that there may be a linear correlation between feed cost, medical service cost, depreciation of fixed assets, labor cost, total gross output value, and net profit in the three farming scales of large, medium, and small. Therefore, in this paper, the Pearson correlation coefficients of different items in the above three categories were first calculated and analyzed using SPSS software. The results show that for free-range dairy farming type, the Pearson correlation coefficient between net profit and total output value is as high as 0.955, with strong positive correlation; the correlation coefficients between net profit and feed expense, depreciation of fixed assets and labor cost are all above 0.8, with strong positive correlation; the absolute value of correlation coefficient between net profit and medical and epidemic prevention expense is only 0.514, with weak negative correlation. The absolute value of the correlation coefficient is only 0.514, and the negative correlation is weak. In addition, the Pearson correlation coefficients between the six variables were above 0.9 for small-, medium- and large-scale farming, with strong positive correlations.

Therefore, for free-range farming, small, medium, and large-scale dairy farming types, an initial attempt can be made to develop a multiple linear regression model between the six variables of feed costs, medical services, depreciation of fixed assets, labor costs, total gross output value and net profit for further investigation.

2.2. Selection of Independent and Dependent Variables

In this study, China's dairy farming income is set as an explanatory variable, affected by both sales revenue and total costs: in terms of dairy farming sales revenue, the main products and by-products sales revenue are included in the calculation, that is, the total output value is included in the explanatory variable; in terms of the total cost of raw milk production, due to the utility costs, technical services, repair and maintenance costs, as well as insurance costs, finance costs, etc. account for a relatively small share of the cost of cost-effectiveness of the main body of the impact of the degree of the farming is not obvious, therefore, in terms of the cost of determining the cost of feed costs, medical prevention, and vaccination costs, depreciation of fixed assets and labor costs are included as an explanatory variable.

2.3. Basic assumptions of the multiple linear regression model

First, it is assumed that the independent variables feed cost, medical service cost, depreciation of fixed assets, labor cost, and total output value combined are linearly independent of the random error term ε . Second, it is assumed that the mathematical expectation of the random error term is 0 and its

variance is a constant; finally, it is assumed that the random error term obeys a normal distribution and that the sequence of the error terms is not subject to serial autocorrelation, i.e., each random perturbation is independent of the other random perturbations.

2.4. Construction of regression models

Let the linear regression model of the explained variable Y with p explanatory variables X_1, X_2, X_p can be expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1)$$

And the above equation is a multiple linear regression model when $p \geq 2$. This equation shows that the dependent variable Y is determined by two parts: one part is a linear function of p independent variables $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$, and the other part is the error term random variable ε , which represents the effect of the rest of the unknowns in the model on Y except for the p independent variables.

Thus, in this study, the explanatory variables were determined based on the analysis of the net profit Y of dairy farming with each influencing factor as feed cost (X_1), medical and epidemic prevention cost (X_2), fixed asset cost (X_3), labor cost (X_4), and total gross output value (X_5). The modeling was done as follows:

2.5. Results and Tests of Regression Models

2.5.1. nomic significance test for coefficients

According to the economic formula of "net profit = total output value - total cost", net profit is in the same direction with the total output value and in the opposite direction with the total cost, so the regression coefficient of the total output value, i.e. β_5 , is theoretically positive, and the regression coefficients of all the independent variables constituting the total cost, i.e. $\beta_1, \beta_2, \beta_3$, and β_4 , are all theoretically negative. Observing the estimation results of the coefficients of the regression model calculated by SPSS, it can be seen that under the premise of net profit as the dependent variable, the multiple linear regression fit of the free-range type and the small-scale farming type fully satisfies the above theoretical situation; in medium- and large-scale farming, although there are positive regression coefficients of the cost items in the first multiple linear regression, the significance level of them is not strong, and the regression coefficients of the independent variables constituting the cost items are all positive after the stepwise regression. The regression coefficients of the independent variables were all negative, which finally satisfied the theoretical situation. In summary, the coefficients of the different types of multiple linear regression models passed the economic significance test and conformed to the actual economic significance.

2.5.2. sts of the regression model as a whole and regression coefficients under the backyard type

The degree of influence of independent variables $X_1 \sim X_5$ on net profit is estimated by the least squares method, and according to the results of the model summary, it can be seen that the regression equation has a good overall goodness-of-fit. The D-W value is close to 2, which indicates that the data series does not have significant autocorrelation, and it meets the conditions of linear regression independence. According to the Table 2, the F-value of the significance test of the regression equation is quite big, and the corresponding P-value is smaller than the significance level $\alpha = 0.05$, indicating that the linear relationship between the variables is significant. According to the estimation of model coefficients, the significance of all five explanatory variables converged infinitely to 0 and the significance of the constant term was smaller than 0.05, therefore, the model coefficients all passed the t-test at the 5% significance level, and there was no multiple covariance check. Therefore, the final model result under free range farming method is:

$$Y = 64.340 - 1.08X_1 - 1.693X_2 - 1.125X_3 - 0.970X_4 + 0.990X_5 \quad (2)$$

2.5.3. sts of the regression model as a whole and regression coefficients under the small scale type

For the type of small-scale dairy farming, the overall decidable coefficient of the model of multiple linear regression is $R^2=1.000$ in Table 1, which shows that the influence of factors other than the independent variable on the dependent variable net profit is very small and can be ignored. In Table 2, the D-W value can be regarded as a non-significant linear relationship between the variables and each other within a certain tolerance range. According to the Table 3, the F-statistic value is much larger than the test statistic of 1.96 at the confidence level of 95%, and the P-value is much smaller than the significance level of α . Therefore, it can be inferred that, at the level of 95% certainty, it can be assumed that the independent variable has a significant effect on the dependent variable. According to the regression coefficient estimation table it can be seen that except for the coefficient of medical epidemic prevention costs β_2 significance of 0.379 is greater than 0.05 did not pass the significance test, the other coefficients have passed the test of significance at the level of $\alpha = 0.05$. Therefore, excluding the independent variable of medical epidemic prevention fee, the model equation under the condition of small-scale dairy farming is:

$$Y = -316.744 - 0.997X_1 - 1.064X_3 - 0.997X_4 + 0.992X_5 \quad (3)$$

2.5.4. sts of the regression model as a whole and regression coefficients under the medium scale type

Under medium-scale farming conditions, $R^2 = 0.996$, indicating that about 99.6% of the dependent variable net profit can be explained by the relationship between X and Y in multiple linear regression. Similarly, the regression model as a whole passed the significance test. However, according to the Table 4, in terms of the tests for the regression coefficients and constant terms, the regression coefficients for the constants, the regression coefficients for the medical and epidemic prevention costs, and the regression coefficients for the labor costs failed the t-test. Moreover, according to the results of covariance statistics, the variance inflation factor VIF is much larger than 10, which indicates that it is very likely that the significance level is problematic due to the multicollinearity among independent variables. Therefore, stepwise regression was performed for the first derived multiple regression model under medium-scale farming conditions. Under the premise of keeping the overall goodness of fit and the overall significance level of the model basically unchanged, the total production value, feed cost, depreciation of fixed assets, and labor cost were added to the model for the significance test of the coefficients, and it was ensured that each newly added independent variable could not affect the t-test of the existing coefficients in the model until the maximum possible combinations of independent variables were obtained to meet the conditions of the t-test. test conditions. At this point, although the variance inflation factor VIF still shows the possibility of correlation between the independent variables, the degree of correlation has been greatly reduced, and the impact on the test of significance has been greatly weakened. Therefore, the regression equation of net profit of medium-scale dairy farming is:

$$Y = -1094.118 - 0.849X_1 - 1.333X_3 - 0.819X_4 + 0.921X_5 \quad (4)$$

2.5.5. ts of the regression model as a whole and regression coefficients under large-scale type

According to the Table 5, Under the type of large-scale dairy farming, the overall fit of the model was good, and the significance level of the F-test was high, but the regression coefficients of the feed cost term, the medical and epidemic prevention cost term and the fixed asset depreciation term did not pass the t-test, so we continued to optimize the model fit by using the stepwise regression method. In the model, the depreciation of fixed assets, the total output value, the labor cost, the depreciation of fixed assets, and the feed cost were added step by step, and the multiple regression equation under the conditions of large-scale dairy farming passed all the tests was finally obtained:

$$Y = -1757.819 - 0.419X_1 - 1.915X_4 + 0.569X_5 \quad (5)$$

Table 1. Model summaries

Model	R	R-square	Adjusted R-square	Error in Standard Estimates	D-W
Backyard	1.000	1.000	1.000	13.993263056898680	2.449
Small	1.000	1.000	1.000	31.924691870584486	2.586
Medium	.998	.996	.994	177.205608304553750	2.632
Large	.996	.992	.990	234.35115	2.120

Table 2. ANOVA(Backyard)

Model	Sum of Squares	df	Mean Square	F	Significance
SSE	41369187.812	5	8273837.562	42254.113	.000
SSR	1762.303	9	195.811		
SST	41370950.115	14			

Table 3. ANOVA(Small)

Model	Sum of Squares	df	Mean Square	F	Significance
SSE	50716092.246	5	10143218.449	9952.275	.000
SSR	9172.674	9	1019.186		
SST	50725264.920	14			

Table 4. ANOVA(Medium)

Model	Sum of Squares	df	Mean Square	F	Significance
SSE	76749719.031	4	19187429.758	611.029	.000
SSR	314018.276	10	31401.828		
SST	77063737.308	14			

Table 5. ANOVA(Large)

Model	Sum of Squares	df	Mean Square	F	Significance
SSE	79529730.049	3	26509910.016	482.696	.000
SSR	604125.086	11	54920.462		
SST	80133855.135	14			

3. Results

The coefficients of explanatory variables under different types of dairy farming in China were obtained from the above model collation, as shown in Table 6. The results of regression model analysis show that under the four types of dairy farming, feed expenses, depreciation of fixed assets and labor costs and sales revenue have a greater degree of influence on the net profit of dairy farming. Among all explanatory variables, only sales revenue and net profit showed a positive correlation, i.e., the increase in sales revenue would improve the efficiency of dairy farming. Feed cost, medical and epidemic prevention cost, fixed asset depreciation cost, and labor cost were negatively related to net profit, indicating that the increase in the expenditure of the above cost items would cause a decrease in net profit.

Table 6. Schedule of coefficients

Type	Feed cost	Medical cost	Depreciation of fixed assets	Labor cost	Total gross output value
Backyard	-1.018	-1.693	-1.125	-0.97	0.99

Small	-0.997		-1.064	-0.997	0.992
Medium	-0.849		-1.333	-0.819	0.921
Large	-0.419			-0.915	0.569

Under the free-range mode, the impacts on dairy farming returns are, in descending order, medical and epidemic prevention costs, depreciation of fixed assets, feed costs, total output value and labor costs. Feed cost has the most significant impact on dairy farming income, with an impact coefficient of about 1.018, indicating that for every unit increase in feed cost input, the net profit of dairy farming will be reduced by about 1.018 percentage points on average, which indirectly indicates that the feed utilization rate of free-range dairy farmers needs to be improved. The impact of medical and epidemic prevention costs on farm income is about 1.693, indicating that for every unit increase in medical and epidemic prevention costs, the net profit of dairy farming will be reduced by 1.693 percentage points on average. The impact of fixed asset depreciation and labor cost on dairy cows is also more significant, with impact coefficients of about -1.125 and -0.97 respectively, i.e., the increase of fixed asset depreciation and labor cost will reduce the profit of dairy farming accordingly. The total output value changes in the same direction as the net profit of dairy farming, and the correlation coefficient is about 0.99, which shows that the increase of the total output value will increase the profit of dairy farming more significantly. Through the analysis of the comprehensive index system above, it is known that the main factors affecting the total output value of dairy farming are the price of the main product and the output of the main product, in other words, the output of fresh milk and its purchase price have a significant impact on the net profit of dairy farming in the form of free-range farming.

Under small-scale farming, the factors that affect net profit in descending order are depreciation of fixed assets, feed and labor costs, and total gross output value. Among them, the depreciation of fixed assets has the most significant impact on the net profit of farming, and for every unit increase in depreciation of fixed assets, the return of dairy farming will be reduced by about 1.064 percentage points on average accordingly. For every unit increase in feed and labor costs, the net profit of farming will be reduced by about 0.997 percentage points on average. In contrast, for every unit increase in gross output, farming returns will increase by an average of about 0.992 percentage points.

Under the medium-scale farming mode, the effects on dairy farming returns are, in descending order, depreciation of fixed assets, gross output value, feed cost and labor cost. Among them, the depreciation of fixed assets has a significant dominant impact on farm income, and for every unit increase in depreciation of fixed assets, the net profit of dairy farming decreases by about 1.333 percentage points on average. For every additional unit of gross production value generated by the farmers, the net profit was about 0.921 percentage points higher on average.

In the large-scale farming mode, the number of factors affecting net profit is significantly lower than in the free-range mode, and the impact of each factor on its farm income is more balanced. Among them, labor cost dominates the change in profitability, with each unit increase in labor cost decreasing the net profit of dairy farming by about 0.915 percentage points on average. The effect of a one-unit increase in both gross production value and feed costs on farm profitability is nearly offset.

Regardless of the type of dairy farming, the depreciation expense of fixed assets has a significant impact on the income of dairy farming. This phenomenon indicates that the degree of mechanization and specialization of dairy farming in China is very high, and the wear and maintenance of production equipment dominates the whole process of dairy farming. In addition, with the expansion of farming scale, the impact of feed cost on net profit gradually decreases, which can be inferred that although the expansion of farming scale increases the absolute demand for feed by farmers, their bargaining power is relatively higher, and bulk purchasing can depress the unit price of feed, and under its combined influence, the impact of feed cost on the net profit of dairy cattle farming gradually decreases. In addition, the effects of labor costs and the size of total production value on returns are largely independent of farm size.

4. Suggestions for promoting cost reduction and income generation

First of all, we should actively promote backyard, small-scale, medium-scale farming main body moderate to large-scale farming transformation and upgrading. The government can reduce the access threshold of a variety of financing channels, introduce certain preferential policies, strengthen scientific and technological innovation and the transformation of the results of elemental inputs [8], the development of a variety of credit preferential types of assistance to farmers to invest in more assets in order to expand the scale of farming; farmers can choose to withdraw from the farming industry, or cows in the form of physical capital into the stock, investment, and other large-scale farming subject to jointly set up a cooperative or set up a joint-stock company to Improve the level of farming scale; Secondly, we should strictly control the total expenditure of farming costs and optimize the composition of cost item expenditure. For material and service costs, straw silage technology should be vigorously promoted, which is conducive to improving the efficiency of dairy farming, and at the same time, to a certain extent, is conducive to environmental protection and improving social benefits, and secondly, we should strictly calculate the amount of scientific feeding of dairy cows as well as the ration of different feed categories, so as to improve the utility of the unit cost of feed; for the expenditure of labor costs, due to the poorer working environment in the farms, the low threshold of job entry, plus the The characteristics of wage rigidity, reduce employee wages instead of reducing production costs, increase efficiency, but may cause the loss of labor; Finally, to actively build the dairy industry benefit sharing market mechanism. First, farm owners should pay close attention to the price of raw milk, improve the speed of response to market changes, and reduce their own risk costs and losses due to information asymmetry; second, the main body of the dairy industry should strengthen the connection with the third party, accelerate the connotative development based on technological progress, and continue to push forward the "grain-to-feed" and crack the short board of feed grain supply, and improve the dairy industry's profitability. Secondly, the main body of farming should strengthen the contact with the third party to accelerate the development based on technological progress, continue to promote the "grain to feed" and crack the feed grain supply, improve the informationization and intelligence of the dairy industry [9]. Third, sales need to make full use of the existing network platform, organizing local youth with skills and knowledge, actively training network sales means, supplementing the existing traditional sales methods, can better broaden the sales channels, and steadily increase the sales volume [10].

5. Conclusion

Dairy farming is an important part of the development of the dairy industry, so the future research on dairy farming in China is a long way to go. According to the rational economic man hypothesis, how to improve the income of dairy farmers is the key to promote the long-term development of dairy farming. Therefore, for the government, dairy farming and related enterprises, it is necessary to establish a long-term communication mechanism, in order to realize the "cost reduction and increase income" basis, and jointly contribute to the development of China's dairy farming, and thus improve the confidence of the national milk drinking, improve the quality of life of the people, and enhance the sense of well-being of the people.

References

- [1] DONG Xiaoxia, LI Mengjiao, LIU Haomiao. Comparative analysis of cost-effectiveness of dairy cattle breeding in advantageous regions of China[J]. China Food and Nutrition,2013,19(12):19-23.
- [2] LANG Yu, WANG Guixia, WU Peirong. Changes and causes of regional layout of dairy farming in China--an empirical study based on national provincial panel data[J]. China Agricultural Resources and Zoning, 2021, 42(1):127-134.
- [3] YUANG Honghua, CHENG Jinhua, WANG Huili. Research on dairy cattle breeding structure and cost-benefit variation in Jiangsu Province[J]. Jiangsu Agricultural Science,2022,50(23):235-241.

- [4] QIAN Guixia, GUO Jianjun. New Issues and Countermeasures for the Development of China's Dairy Industry and Future Trends[J]. *Agricultural Outlook*,2008(04):17-21.
- [5] [Ogunyinka, EO., Ajibefun, L. A. Determinants of Technical Efficiency among Dairy Farms in Wisconsin[J].*Journal of Dairy Science*, 2010(93).
- [6] Feng Dongdong,Wang Juan,Zhu Jiajun et al. Development of dairy farming environment monitoring system based on Internet of things[J]. *Journal of Hebei Agricultural University*,2023,46(01):2023.0015.
- [7] Zhang Xing. Analysis of development status and countermeasures of dairy cattle breeding industry. *Modern Animal Husbandry Science and Technology* .01(2023):25-28.
- [8] Li Junru,Shi Zizhong,Hu Xiangdong. Characterization of spatial and temporal variation of total factor productivity in China's dairy industry[J]. *China Agricultural Resources and Zoning*,2022,43(7):92-103. DOI:10.7621/cjarrp.1005-9121.20220710.
- [9] SHI Yu, YAO Peixin. Cost-benefit analysis of dairy farming in different scales in China[J]. *China Dairy Industry*,2021,49(10):44-48. DOI:10.19827/j.issn1001-2230.2021.10.009.
- [10] Liu P. High-quality characteristic farming of dairy cattle in Shaanxi region to promote rural revitalization[J]. *Rural Science Experiment*,2023(3):154-156.