Special issue of the 9th International Conference held on 6/6/2021 , Under the Slogan (Sustainable Development through Qualitative Education and Digital Revolution)

Study the social and economic factors that affects the adoption of management best practices in livestock farming: An applied study in ECER,Malaysia

prof.Dr. Fayyadh Abdulla Ali
Wasit university
College of Administration and Economics

Azzam bdullahA Tawfeeq
Baghdad College of
Economic Science University

Abstract

Smallholder livestock farmers in Malaysia, in public, and in East Coast Economic Region (ECER), specially, have trouble in cattle production in word of adopt Best Management Practices(BMPs) in cattle farming.

The importance of this research lies in bridging the knowledge gap between farmers 'adoption of BMP systems and policy-making in the livestock industry, it is the basic understanding of the socio-economic variables of farmers that will help policy makers understand farmers 'requirements to improve their adoption on BMPs and then improve livestock production.

The main aim of this research is to build a logistic regression model between adoption of BMPs and social--economic factors on smallholder livestock farmer in ECER, Malaysia

The research was divided into two parts, the first part dealt with the importance and goal of the research in addition to the theoretical aspect of the logistical model, while the second part included how to apply the logistical model in terms of identifying the variables affecting the adoption of best management practices(BMPs) in livestock farming.

الملخص

يواجه مربو الماشية أصحاب الحيازات الصغيرة في ماليزيا ، عموما، وفي المنطقة الاقتصادية للساحل الشرقي (ECER) على وجه الخصوص ، مشاكل في إنتاج الماشية نتيجة تبني أفضل ممارسات الإدارة (BMPs) في تربية الماشية.

تكمن أهمية هذا البحث في سد الفجوة المعرفية بين تبني المزارعين لأنظمة BMP وصنع السياسات في صناعة الثروة الحيوانية، فإن الفهم الأساسي للمتغيرات الاجتماعية والاقتصادية للمزارعين هو الذي سيساعد صانعي السياسات على فهم متطلبات المزارعين إلى تحسين اعتمادها على BMPs ومن ثم تحسين الإنتاج الحيواني.

الهدف الرئيسي من هذا البحث هو بناء نموذج انحدار لوجستي بين اعتماد BMPs والعوامل الاجتماعية والاقتصادية على صغار المزارعين في الماشية في ECER ، ماليزيا.

تم تقسيم البحث إلى قسمين ، تناول الجزء الأول أهمية البحث وهدفه بالإضافة إلى الجانب النظري للنموذج اللوجستي ، بينما تضمن الجزء الثاني كيفية تطبيق النموذج اللوجستي من حيث تحديد المتغيرات المؤثرة على البحث. تبني أفضل ممارسات الإدارة في تربية الماشية.

1.Introduction

Smallholder livestock farmers in Malaysia, in public, and in East Coast Economic Region (ECER), specially, have trouble in cattle production in word of adopt Best Management Practices(BMPs) in cattle farming. these aspects are of essential importance because of having some direct or indirect leverage on industry finally, chiefly on production.

One of definition of BMPs[1] as activity, process, technique or method , that is in general understood to become more influential in accomplish a certain result or achieve a task than other classical activity, process, technique or methods.

From this, in this research , BMPs are defined as the techniques which are more dynamic in animal feeding, breeding and healthcare to fulfill the best result of animal production.[5]

The BMPs have be agriculture program for so long in most all polities in the world. These are important for various seasons such as:[5]

i. Control of nutrient loss
ii. Getting cost advantage
iii. Improving the environmental performance of livestock sector and
iv. Its impact on water quality and increasing efficiency and profitability.

BMPs have specified important in livestock production, among others, to get best genetic, preferable calves, better overall animal health with regard to present farm environment, and better overall animal welfare.

livestock farmers with various socia-economic profiles could perform a significant role in improving livestock production by adopting of cattle (BMPs)

in addition , the socia-economic factor is the most significant , that may impact the adoption of BMPs in livestock farming in ECER states.

2.Significance of research

The results of this research would be significant of datum to bridge knowledge gap between farmers' adoption of BMPs and politics making in livestock industry.

It is the basal understanding of farmer's social and economic variables that will assist politics makers to understand the requirements of farmers to improve in adoption of BMPs and then to improve livestock production**.**

3. Search Problem

livestock farming in Malaysia are yet backward behind comparing to other agriculture sectors. The equipping of livestock products are not growing sufficiently. Where only the production of 30% of beef and 9% of milk to meet the local need, while the rest is imported from overseas. Adopting the BMPs in livestock is one of the most important factors for improving livestock industry in Malaysia. The socio-economic variables of farmers may effect this adoption. sufficient information on this issue, especially in ECER of Malaysia is not available. Therefore, this paper focuses on the impact of social and economic variables of smallholder farmers on the adoption of best management practices in the ECER ,Malaysia.

4. Objective of research

The main aim of this research is to build a logistic regression model between adoption of BMPs and social--economic factors on smallholder livestock farmer in ECER, ’MLAYSIA

5.Scope of research

Though BMPs can lid a large spectrum of environmental, physical, economic and social aspects it is significant to determine them to a more manageable research. for this cause, only BMPs in feeding, breeding, and healthcare have been selected as research subject, further more, these aspects of BMPs have been determined as Malaysian government's aim in the area of agricultural administration for livestock farming development.

the research was made in ECER , Malaysia, that centered on the determine of socia-economic variables of smallholder livestock farmers and their level of adoption of BMPs in livestock farming[5]

6.Logistic Regression

Although regression analysis achieves most of the objectives of scientific research and its methods are the basic part of any data analysis aimed at studying and interpreting the relationship between the dependent variable and the independent variables, it fails to describe and interpret the relationships between the explanatory variables and the dependent variable if the latter is binary, The nature of the dependent variable is a continuous, not a taxonomic, taking into consideration the prevalence of this type of dependent variable in the case of the study of phenomena. Hence the need to develop new statistical methods that have the power of linear regression in arriving at the best possible equations and simultaneously address the problem of the inability to apply models of linear regression analysis in the case of binary dependent variables value

In Linear Regression our interest is only in the case of the dependent variable of the continuous type, but there are cases in which the dependent variable is of the binary type. In this case, we use the logistic regression model.Therefore, the logistic regression is only a type of regression where the dependent variable is a qualitative variable that may take two values (the binary logistic regression) and may take more than two values (multiple logistic regression). However, the logistic regression is an objective not to explain the change in the values of the dependent variable, but it explain the probability of occurrence and non-occurrence of the phenomenon under study, and represents the logistic regression equation as follows:

$$P_i = E(X_i) = \frac{e^{(B_0+B_1x_1+\cdots+B_kx_k)}}{1+e^{(B_0+B_1x_1+\cdots+B_kx_k)}} \;\dots\dots(1)$$

P_i represents the probability of the class or attribute i of the dependent variable.

B represent the parameters of the model.

Y: represent dependent variable.

X: represent independent variables.

The equation shows that the relationship between the dependent variable and the independent variables is nonlinear, so there are several transfers that can be made to make the relationship linear. One of the most important of these conversions is the logit conversion.

$$L = ln\left(\frac{P_i}{1-P_i}\right) = B_0 + B_1x_1 + \cdots + B_kx_k \;\dots\dots(2)$$

$\frac{P_i}{1-P_i}$ represent the odd ratio, which is the cornerstone of interpreting the parameters of the model, is explained by the type of explanatory variable.

The interpretation of regression parameters in the logistic regression model is similar to the classical regression model, however, rather than describing the change in the expected value of Y, they now describe the change in the probability of viewing Y = 1 shown on the logarithmic scale. [10]

7. The concept of logistic regression

Logistics regression analysis mainly focuses on the classification of individuals in different groups. The logistic regression analysis is also called the Binary Logistic Regression Analysis, the Multinominal Logistic Regression Analysis, and the Ordinal Logistic Regression Analysis [2] Depending on the type of scale where the dependent variable is measured and the number of dependent variable categories is divided into two parts: "single-variable logistic regression" and "multiple logistic regression" and independent variables can be quantitative or qualitative as well.

However, logistic regression analysis differs from the Multiple regression In some points, logistic regression analysis, unlike multiple regression analysis, does not require assumptions to satisfy with respect to the distribution of independent variables. Therefore, logistic regression analysis is much more flexible than the Multiple regression, it is logical to say that it is easier to interpret the mathematical model obtained as a result of analysis by logistic regression analysis [9]

8. Assumptions of logistic regression[11]

Logistic regression does not assume a linear relationship between dependent variables and independent.

- The dependent variable must be a qualitative variable.
- Independent variables do not need to be a period, nor are they distributed normally, nor linearly linked, and equal to the variation within each group.
- Categories (groups) must be mutually independent
- Larger samples than linear regression are needed because larger potential coefficients are large sample estimates. It is recommended that at least 50 cases per forecast.

9. Estimating of model parameters

Instead of finding the best matching line by minimizing error squaring as in the Ordinary Least Square(OLS) method, we use a different approach with logistic regression, the maximum Likelihood[3]. To find the smallest possible deviation between observed and estimated values in this way, the computer performs a number of iterations We obtain a number of different solutions so that we get the least deviation or better matching that gives the final value of the deviation, which is commonly referred to as "negative two log likelihood" (in the program spss referred to as "-2 Log Likelihood").

The deviation statistic is called -2LL by Cohen et al. (2003) and Pedazur and called D by other authors (Hosmer, Lemeshow, & Sturdivant, 2013) which can be viewed as a value for χ^2

10. Evaluate the model

The final stage of building the model is the evaluation of the model which is done in two stages:

10.1. First stage: Overall evaluation of the model

This is done through the quality standards of conciliation, and the overall morphological tests of the model.

- **Quality Assurance Standards**

Is a statistical standard that measures the explanatory power of the model. These criteria are considered as alternatives to coefficient of determination of linear regression. These are called semantic coefficients of determination, which are calculated by comparing the force of the model interpretation without explanatory variables and then after the introduction of explanatory variables. The most important of these standards is Cox & Snell, which was developed by Nagelkerke [8] and is calculated as follows:

$$R^2{}_{Cox\ \&\ Snell} = 1 - \left(\frac{-2LL_{null}}{-2LL_k}\right)^{\frac{2}{n}} \ldots\ldots(3)$$

$-2LL_{null}$ represents the likelihood function in case the model contains only the constant.

$-2LL_k$ the likelihood function in the case of the model contains all explanatory variables.

n represents the sample size

Because the value of R^2 can not reach 1.0, the Nagelkerke has modified it. This patch has increased R^2 to Cox & Snell to make R^2 value up to 1.0.

$$R^2{}_{Nagelkerke} = \frac{1-\left(\frac{-2LL_{null}}{-2LL_k}\right)^{\frac{2}{n}}}{1-(-2LL_{null})^{\frac{2}{n}}} \quad \ldots\ldots(4)$$

The values R^2 of Cox & Snell and R^2 of Nagelkerke represent two different ways of predicting the interpreted variance in the dependent variable in the same R^2 in multiple regression [4].

- **Test the significant of the model**

The purpose of this test is to know the overall mean of the model. In the sense, are all the parameters of the interpreted variables equal to zero or there is at least one parameter equal to zero, using the same principle in the case of the normal regression, which is the comparison of the expected values in the case of the model without the variables explained with the expected values in the model Containing independent variables, this test is called Hosmer And Lemeshow [6]

This test is used to see if the model represents the data well or not. The Chi- square test is used for conformance to assess the difference between observable and predicted values, and test the following hypotheses

H_0 equals the observed values with predicted values, that is, the model represents the data well

H_1 The observed values are not equal to predicted values, meaning that the model does not represent the data well

The decision to accept the null hypothesis shall be determined if the probability value of the χ^2 statistic is greater than the level specified by the researcher.

χ^2 is used as a measure of the goodness of fit of the observed and predicted values. Here we use it as a measure to reconcile the model to match the observed values (Y) of the predicted values $\hat{Y}$

The greatest difference of observed values from predicted values gives the worst match to the model and the objective always minimizes differences [7]

10.2. Second stage : Tests significant of parameters

We conduct these tests to determine the statistical significance of each variable separately, and the wald test is usually used

- **Wald Test** [12]

The wald test is one of the most important tests used to judge the effect of the explanatory variable on the dependent variable. It is calculated from the model variables to determine the variable to be deleted. If the variable is not qualitative, the wald statistic is defined as:

$$wald_i = \frac{\hat{\beta}_i^2}{\hat{\sigma}_{\hat{\beta}_i}^2} \ldots\ldots(5)$$

If the variable is qualitative, the wald statistic is calculated as follows:

Assume $\hat{\beta}_i$ vector estimates of maximum likelihood associated with m-1 dummy variables

The C is a common approximation matrix for $\hat{\beta}_i$. The wald statistic is:

$$wald_i = \hat{\beta}_i{}' c^{-1} \hat{\beta}_i \ldots.(6)$$

The approximation of the wald statistic is χ^2 with a degree of freedom equal to the number of estimated parameters.

11. Application

11.1. sampling and sample

The Stratified random sampling was used to collect data from smallholder livestock farmers of ECER, Malaysia. Three villages were selected from each district of ECER states . generally ,33 villages from various district of ECER were selected according to the density of number of farmers. The sample size was 607 farmers, 97 from Pahang ,100 from Terengganu, 110 from eastern part of Johor, and 300 from Kelentana. [5]

11.2.Resrearch variables[5]

11.2.1. dependent variable

- level of adoption: This variable represents the dependent variable and its value is determined based on 30 BMPs and these thirty BMPs have associated with health care, feeding and breeding and takes two values zero or one (it takes the value zero if the adoption level is less than 10 and one if it is greater or equal to 10)

11.2.2.Independent variables (13 variables)

- **Farm size(acres):** The variable takes two values 1 if the leased or owned area is less or equal 3, but if it is greater than 3 then it takes the value 2.
- **Farmer annual income(RM):** This variable represents the annual income that the farmer gets as a result of selling his animal products and takes two values 1 if the income is less or equal to 9000RM and takes the value 2 other wise.
- **Farmer's age(year):** This variable represents the age of the farmer (in year) and takes the value 1 if the age is less than or equal 35 and 2 if the age is between 35-60 and 3 otherwise
- **Farmer's education:** This variable represents the academic achievement of the farmer and takes the values 0 in the case of the farmer who does not read or write , 1 if he reads and writes, 2 if he holds a primary certificate, 3 if he holds a middle school diploma and 4 if he holds a bachelor's degree.
- **Number of cattle:** This variable represents the number of livestock the farmer owns, and the variable takes the value 1 if the number of livestock he owns is less than 12 and 2 if he is more than 12.
- **Current knowledge in reproductive management:** This variable represents the current knowledge of reproduction and takes the value 1 if his knowledge is weak, 2 if it is medium and 3 if it is high.
- **Intention for training course:** This variable represents the intention to train and takes the values 1 if his intention is weak and 2 in the case it is medium and 3 if the intention is high.
- **Cattle farming as a main occupation:** This variable represents whether the main occupation is livestock and takes the value 1 if the main occupation of farmer and it is zero and otherwise
- **Sharing of training course:** This variable represents the participation of the farmer in the training courses and takes the value 1 in the case of participation and zero in the case of non-participation.
- **Family size(persons):** This variable represents the number of family members.

- Family members working in farm-Taking care of animal (persons)

- Taking care of animal: It is classified into two classes, the first category takes the value 2 if the farmer takes care of himself and the value 1 if he takes care of another person

- Type of owned cattle: This variable indicates the type of livestock owned and has been divided into four categories 1 for local breed,2 for crossbred,3 for improved breed and 4 for combined breed.

11.3. statistical analysis

- The statistical program (SPSS ver 17.0) was used in the statistical analysis as follows:

First: The overall evaluation of the model

In order to know the explanatory power of the model, the results of the embedded model are often compared only to the constant term (without explanatory variables) with the embodiment of the explanatory variables.

Table(1) Iterations and the value of the probability function of the model without explanatory variables

Iteration		-2 Log likelihood	Coefficients
			Constant
Step 0	1	818.003	-.392-
	2	817.999	-.397-
	3	817.999	-.397-

Table (1) represents the value of the likelihood function in the case of the model that contains the fixed term only, that is, without explanatory variables. The method of estimating the parameters of the logistic regression model is a iterative method, meaning that the calculation process is repeated until the parameter values converge at a specific permissible error and at the smallest value of the possibility function where the stop is observed at the third attempt, where the value of the function reached 817.999, which is equal to the previous attempt this value we will compare with The embodiment of the explanatory variables If the introduction of the explanatory variables leads to a decrease in this value, it can be said that the explanatory variables affect the dependent variable.

Table (2) estimation of the constant parameter only for the logistic regression model

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-.397-	.083	23.025	1	.000	.672

Table (2) shows that the value of the constant -0.397 and that the wald statistic is 23.025 which is highly significant less than 0.001, so we reject the null and accept the alternative hypothesis, i.e. accept the significance of the constant.

Table (3) iterations and value of the likelihood function of the model including the explained variables

iteration		-2 Log likelihood	Coefficients													
			C	Farm size	Farmr incom	Farmr age	Farmer Educ. level	Num. Of cattle	Curr. Know.	Int. for train.	Catt Main Occup.	Sharing Train cour	Family size	Family Mem.	Take Care animal	Type Own breed
Step 1	1	723.748	-1.401-	-.287-	.552	-.084-	.358	-.339-	.498	.098	-.167-	.376	.054	-.121-	-.471-	.163
	2	721.578	-1.614-	-.364-	.661	-.097-	.418	-.397-	.557	.115	-.188-	.436	.066	-.146-	-.553-	.203
	3	721.571	-1.627-	-.370-	.667	-.098-	.421	-.400-	.560	.116	-.189-	.439	.067	-.148-	-.558-	.205
	4	721.571	-1.627-	-.370-	.667	-.098-	.421	-.400-	.561	.116	-.189-	.439	.067	-.148-	-.558-	.205

Table (3) shows the following:

First: The stability of the model parameters at the forth step, where we note that the values of the third step are equal to the values of the forth step.

Second: The estimation process stopped at the smallest value of the likelihood function (721.571), and we note that this value is much smaller than the value of the likelihood function in the case of the model that includes the fixed term only (817.999) and this indicates a relationship between the dependent variable and the explained variables.

Table (4) criteria of goodness of fit

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	721.772	.149	.201

Table (4) represents the semi coefficient of determation that express the explanatory power of the model, where the value of $R^2_{Cox\,\&\,Snell}$ (0.149) while the value of $R^2_{Nagelkerke}$ (0.201) means that% 20.1 of the variance in the dependent variable is explained by the explained variables.

Table (5) The significance test for the overall parameters of the model (Hosmer-Lemeshow)

Step	Chi-square	df	Sig.
1	5.782	8	.672

Through Table (5), we note that the calculated level of significance (0.672) is greater than level of significance 0.05, and therefore we accept the null hypothesis at the level of significance of 5%, that is, the model fits the data well, which indicates the significance of the parameters of the model in general.

Table(6) Classification Table

	Observed		Predicted		
			adoption		Percentage Correct
			0	1	
Step 1	adoption	0	288	73	79.8
		1	123	123	50.0
	Overall Percentage				67.7

Table (6) shows the correct classification rate for both adoptions and non-adoptions and the overall percentage as we note that the rating for non-adoption is 79.8% higher than adoption 50%, and the overall percentage indicates the model's ability to rate or predict 67.7%, i.e. the error rate is equal to 32.3%.

Table (7) Results of estimating model parameters and odds ratios

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	Farm size	-.404-	.199	4.142	1	.042	.668
	Farmers income	.718	.266	7.266	1	.007	2.050
	Farmer age	-.134-	.144	.860	1	.354	.875
	Farmer education level	.394	.143	7.597	1	.006	1.483
	Number of cattle	-.432-	.273	2.504	1	.114	.649
	current knowledge	.547	.120	20.607	1	.000	1.728
	Intention for training course	.144	.109	1.764	1	.184	1.155
	Cattle main occupation	-.153-	.197	.604	1	.437	.858
	Sharing train course	.436	.214	4.176	1	.041	1.547
	Family size	.068	.034	3.921	1	.048	1.070
	Family members	-.124-	.137	.810	1	.368	.884
	Take care of animal	-.512-	.241	4.515	1	.034	.599
	Type own breeding	.208	.080	6.674	1	.010	1.231
	Constant	-1.641-	.830	3.910	1	.048	.194

The wald test in Table (7) above shows the following:

- The significant of Current knowledge in reproductive management, Farmer education level, Farmers income, Type own breeding, Take care of animal, Sharing train course, Farm size, Family size and constant at the 5% level of significance.
- The remaining variables are not significant, which means that these variables are not important in explaining the level of adoption.

-We note that there is an inverse relationship between the large size of the farm and the level of adoption where the last column indicates that the value is less than one and the same for take care of animal. While we note that there is a positive relationship between the level of adoption and the variables (Current knowledge in reproductive management , Farmer education level, Farmers income, Type own breeding, Sharing train course, Family size) Where the value is greater than one For example, the high incomes of farmers are more likely to adopt twice(2.050) than those with low incomes, and this is reflected in the last column(Exp(B))

12. Conclusions

- The use of the OLS method in estimating the regression model that includes specific dependent variables leads to inaccurate results because some of the estimated values of the dependent variable are greater than (1 or smaller than 0) which are meaningless because the original values of the dependent variable take the two values.
- The results obtained indicate that the most influencing variables in adoption are Current knowledge in reproductive management, Farmer education level, Farmers income, Type own breeding, Take care of animal, Sharing train course, Farm size and Family size.

13.Recommendations

Based on the above, and in light of the study, the researcher proposes the following recommendations:

1. Extending the use of logistic regression in social and economic studies, where its previous uses were limited to the fields of medical and educational sciences**.**
2. Similar studies can be carried out in the Iraqi environment by adopting new administrative policies to develop work in the agricultural sector and meet the local need of most animal products.

References

[1] Anonymous.(2016)Wild Predator loss prevention best management practices for cattle: A guide for cattle producers on how to minimize predation of cattle. Investment Agriculture Federation of British Columbia, Canada.

[2] *ÇOKLUK*, *Ömay*," Logistic Regression: Concept and Application", (2010) Kuram ve Uygulamada

Eğitim Bilimleri / Educational Sciences: Theory & Practice , 1397-1407

[3] Czepiel, Scott A., Maximum Likelihood Estimation of Logistic Regression models: Theory and Implementation, http://czep.net/contact.html

[4] Field, A. (2005). Discovering statistics using SPSS (2nd ed.). London: Sage .

[5] Hassan, Amal Najim,(2017),Socia-Economic factors influencing Adaption of Best Management Practices among cattle farming farmers in East Coast Economic Region(ECER),Malaysia.

[6] Hosmer, David w. and Stanley Lemeshow (2000), Applied Logistic Regression ,2nd Edition, John Wiley and Sons, Inc.

[7] Newsom 1 Data Analysis II Fall (2015), Logistic Regression, Overview: Logistic and OLS Regression Compared.

[8]Newsom 1 Data Analysis II Fall (2010) Multiple Logistic Regression and Model Fit

[9] Poulsen, J., & French, A. (2008). *Discriminant function analysis.* http://userwww.sfsu. edu/~efc/classes /biol710/discrim/discrim.pdf adresinden 22 Kasım 2008 tarihinde edinilmiştir

[10] www.isqr.uni-freiburg.de/**logistic**.pdf

[11] https://www.researchgate.net/file.PostFileLoader.html?id...assetKey

[12] www.upch.edu.pe/facien/fc/dbmbqf/zimic/.../**logistic_regression.pdf**