

CHAPTER 4

DATA, DESCRIPTIVE STATISTICS AND ESTIMATION METHODS

The farm size efficiency of rice based multiple cropping systems is the main interest of this study. As of this, the study area is chosen to be Nong Han subdistrict, San Sai district of Chiang Mai province. This is based on the fact that, in this subdistrict and district, rice comprises over 95% of the rainy season lowland cropping acreage. This means the operation size of rice crop is approximately the same as farm size. In addition, the dry season crops are diversified, farmers grow soybean, dry season rice, potato, tomato, garlic and other crops. This enables us to investigate farm size efficiency in different cropping systems.

This chapter describes the study area and data collection and aggregation. The descriptive statistics of the household and explanatory variables are also presented. Finally, estimation models and methods are discussed.

4.1. The Study Area

San Sai is one of the districts of Chiang Mai province. It covers an area of 198,758 rai, among which 78,412 rai is used for agriculture. There are 19,413 households in the district, with a population of 77,976 people in 1989 (Agricultural Office of San Sai district, 1990).

Nong Han is one of 12 subdistricts in San Sai. It covers an area of 23,248 rai, with 1,673 agricultural households, both are the biggest among the 12 subdistricts of San Sai district. The area is relatively well irrigated,

though there is not enough water in the dry season. The soil is San Sai series, which is a fertile soil in the Chiang Mai valley, and the farm field is relatively homogeneous across farms. Table 4.1. and 4.2. summarize some basic information of San Sai district and Nong Han subdistrict.

Table 4.1. Some Information of San Sai District and Nong Han Subdistrict (1989)

	Total land (rai)	Agr. Land (rai)	Total household	population	Agr. household
San Sai	198,758	79,412	19,413	77,958	11,243
Nong Han	23,248	9,220	1,845	7,798	1,873

Source: Agricultural Office of San Sai District, 1990.

Table 4.2. Major crops in San Sai and Nong Han in 1989 (rai)

	San Sai	Nong Han
Rainy season N. G. rice	10,178	650
Rainy season G. rice	30,091	4,268
Dry season N.G. Rice	21,051	-
Dry season G. rice	1,651	49
Rainy season Soybean	778	20
Dry season soybean	20,849	1,500
Garlic	1,438	571
Tomato	1,899	194
Potato	2,600	1,640

Source: Agricultural Office of San Sai District, 1990.

N.G.: Non Glutinous

G.: Glutinous

Crop production in the dry season is well regionalized. On the subdistrict level, garlic, potato and tomato are grown in only four subdistricts, although soybean is grown in all of the 12 subdistricts. In Nong Han district, soybean is mainly grown in villages No. 1, 2, 3, 7 and 11; potato is grown in villages No. 8 and 4; tomato is grown in villages No. 3, 2, 11 and 5. It is a usual practice that only one dry season crop is grown for each household, which means although cropping systems are complicated in the whole area, each individual farmer has a simple cropping system.

4.2. Data Collection

The study attempts to analyze farm size efficiency of four cropping systems, rice-soybean, rice-potato, rice-tomato, and rice-tomato. A structured questionnaire is designed for each system, and a total of 300 farmers are interviewed, 75 from each of the four cropping systems. These 75 farmers are divided into three groups, 25 farmers each, to represent the small, medium and large farmers in the sample. Within these 3 groups, farmers are selected randomly.

Because of incomplete data and other problems, only 70, 60, 71 and 60 questionnaires from soybean, potato, tomato and garlic farmers, respectively, are good for the final analysis. As will be discussed later, among the 71 observations of tomato growers, only four apply chemical fertilizer 16-20-0, these are omitted in the final analysis. In addition, some farmers only grow crops in the dry season; they do not grow rice in the rainy season. This together with the incomplete data on the rice crop itself and flooding and other problems, leaves only 193 observations for final analysis in the rice model.

Farmers are asked about their plans with regard to farm size, six farmers in the sample report they want to increase farm size, three of them say because they have idle labor in the family, and the other three say because farming gives them good income. Five farmers in the sample want to reduce farm size, two of them have better jobs and another three feel that there are labor demand conflicts between harvesting rice and planting soybeans.

4.3. General Description of the Sample Farmers

4.3.1. Rental Status of the Sample Farmers

Among 193 observed rice farmers (table 4.3.), 106 own their land, and 80 farmers are landless, of whom 62 farmers rent land from landlords and pay a fixed rent of 200 kg rice per year, and another 14 farmers pay 50% of their harvested rice. There are 11 farmers who own part and rent part of their operated land, with either fixed or share rent payment. In all cases of renting, farmers only pay rent for rice, they can farm their land in the dry season with no rent payment.

However, those who do not grow rice in the rainy season but rent some land in the dry season to grow soybean, potato, tomato or garlic are required to pay 500 baht in cash per rai for rent. In the sample, 2, 1, 33 and 25 farmers of soybean, potato, tomato and garlic, respectively, are in this case.

ลิขสิทธิ์ © by Chiang Mai University
All rights reserved

Table 4.3. Rental Status and Cropping Systems of 193 Observations of Rice Farmers

	Soybean	Potato	Tomato	Garlic	Total
Owner Operator	23	30	29	24	106
Fixed Rent	42	5	9	6	62
Share	0	14	0	0	14
Mixed	3	3	0	5	11
Total	68	52	38	35	193

4.3.2. Agricultural Machines and Equipments

Almost all farmers possess such simple equipment as sprayers, hoes and knives. There are 42 farmers who also have small tractors. Since these tractors are also used to plow land for other farmers, thereby earning cash income, machines are not included in the analysis model. It costs about 250 baht per rai for plowing rice, potato, tomato or garlic fields. Therefore the imputed cost is 250 baht per rai for tractor owners as for those with no tractors.

4.4. Explanatory Variables

4.4.1. Operation Size

The descriptive statistics of cropping size distribution is presented in tables 4.4. and 4.5. Rice has the largest average size and is also the most varied across farms. Average operation size is 5.81 rai, with smallest and biggest being 1.5 and 55 rai. Most farmers grow 3-9 rai of rice. Only 11 farmers grow rice on more than 13 rai. Since farmers grow rice on almost all of their lowland

fields in the rainy season, this variable can represent land holding.

Soybean is the largest dry season crop in the sample. It averages 4.77 rai in operation size. The smallest and largest are 1 and 13 rai respectively. However, the operation sizes of potato, tomato and garlic are all small, with averages of 1.99, 2.14 and 2.12 rai respectively. The variation in operation sizes across farms is also small.

Table 4.4. Descriptive Statistics of Cropping Size (rai)

Crop	Mean	Std. Dev.	Coeff. Dev.	Minimum	Maximum
Rice	5.81	4.623	78.22	1.50	55.0
Soybean	4.77	2.428	50.90	1.00	13.0
Potato	1.99	1.301	65.38	.50	9.0
Tomato	2.14	1.058	49.35	1.00	5.0
Garlic	2.12	1.055	49.76	1.00	5.0

4.4.2. Chemical Fertilizer

Three types of chemical fertilizer are applied among the 193 observations of the rice crop, these are urea (46-0-0), 16-20-0 and 21-0-0. A total of 47 farmers applied no fertilizer, 141 farmers applied 16-20-0, and 25 applied either urea or 21-0-0. Fertilizer data of rice are aggregated into two explanatory variables, Pure N and 16-20-0 in kilograms of commercial form. Pure N only captures the nitrogen which comes from urea or 21-0-0.

Table 4.5. Number of farmers in each size group

Rice								
Size (rai)	<3	3-5	5-7	7-9	9-11	11-17	≥17	total
No. farmers	20	63	51	27	21	10	1	193
Percent	10.4	32.6	26.4	14.0	11.9	2.6	0.5	100

Soybean							
Size (rai)	<3	3-5	5-7	7-9	9-11	≥11	Total
No. farmers	13	26	17	9	3	2	70
Percent	18.6	37.1	24.3	12.9	4.3	2.9	100

Potato						
Size (rai)	<1	1--2	2--3	3--4	≥ 4	Total
No. farmers	4	21	27	5	3	60
Percent	6.7	35.0	45.0	8.3	5.0	100.0

Tomato						
Size (rai)	<2	2-3	3-4	4-5	≥ 5	total
No. Farmers	20	33	11	3	4	71
Percent	28.2	46.5	15.5	4.2	5.6	100

Garlic						
Size (rai)	<2	2-3	3-4	4-5	≥ 5 rai	Total
No. Farmers	21	19	15	3	2	60
Percent	35.0	31.7	25.0	5.0	3.33	100

Average rate of pure N and 16-20-0 are 0.59 and 17.5 kg per rai. Pure N and 16-20-0 are coded as FERT1 and FERT2, respectively.

For soybean, farmers apply two formulas of chemical fertilizers. One is 11-8-8, a liquid that is sprayed onto soybean leaves at the stage of soybean flowering and pod formation. Fifty two farmers report applying 11-8-8, with average rate of 1.29 kg per rai among these 52 farmers. Some farmers spray as many as three to four times in the same field. Twenty five farmers apply 2 kg or more per rai, the maximum rate of application is 2.5 kg per rai. In addition, there are eight farmers who apply 12-24-12, with average rate of 24.8 kg per rai among the users. The maximum rate is 43.75 kg per rai.

Due to the different nature of these two types of fertilizer and application methods, they enter the production model separately. 11-8-8 is coded as FERT1 and 12-24-12 is coded as FERT2.

All the observed potato growers use compound fertilizer 12-24-12, with the average application rate 176.7 kg per rai, and ranging from 27.8 to 400 kg per rai. This variable is coded as FERT1.

All observations of tomato growers apply chemical fertilizer. Four formulas of fertilizer are applied, 13-13-21, 15-15-15, urea and 16-20-0. All sampled farmers use urea, with average 26.0 kg per rai. Since there are only four farmers who apply 16-20-0, these observations are omitted. Since chemical fertilizers 13-13-21 and 15-15-15 are similar in their nutrition contents, farmers tend to choose one of them. Fifty one farmers use 15-15-15 but not

13-13-21, nine farmers use 13-13-21 but not 15-15-15, and eleven farmers use both of them. Average rates of 13-13-21 and 15-15-15 application are 9.1 and 26.0 kg per rai, respectively. 13-13-21, 15-15-15 and urea of tomato production are coded as FERT1, FERT2, FERT3, respectively.

Two formulas of fertilizer are observed among the sampled garlic farmers, 13-13-21 and 15-15-15. 17 farmers use both of them, 17 farmers use only 13-13-21 and 26 farmers only use 15-15-15. The result of these two fertilizers is, one kg of pure N is associated with one kg of P_2O_5 . This means N and P have a perfect correlation. Therefore, N and P are added up, and coded as FERT1, K_2O is coded as FERT2.

4.4.3. Labor

Labor use includes all labor inputs from land preparation to harvesting, and is measured in man-days (8 hours = 1 day). When necessary, farmers hire labor from their neighbors, and some farmers exchange their family labor with their neighbors. All types of hired labor, exchange labor, male and female labor are assumed to have the same quality, and are therefore added without weighing.

It must be noted that labor accompanying tractor services is not included in the labor cost for rice, potato, tomato and garlic, since these farmers have to pay 250 baht per rai for such services.

The general statistics of labor inputs for each crop is presented in table 4.6.

Table 4.6. Labor Inputs in Five Crops (man-days /rai)

Crops	Mean	Std. Dev.	Coeff. Dev	Minimum	Maximum
Rice	22.844	5.0730	21.89	9.2	37.0
Soybean	22.368	4.8732	21.79	8.3	35.0
Potato	36.631	5.1672	14.11	8.5	54.0
Tomato	66.363	11.170	16.83	46.0	91.0
Garlic	49.503	13.521	27.31	24.0	66.0

4.4.4. Pesticide and Herbicide

The measurement of these inputs are a little bit cumbersome. Farmers use as many as 5 or more types of pesticides or herbicides on each of their crops, some in liquid, some in powder. Besides, some farmers are not able to remember the names and quantities they apply, they only give the expenditure on pesticides and herbicides. This leads to a difficulty in aggregation.

However, in this study, it is found that pesticide and herbicide prices are rather uniform across farmers. Therefore, one possible aggregation is to use the value of expenditure on pesticide and herbicide as a proxy. This kind of aggregation was also used by Heady and Dillon (1961), Wiboonpongse(1983) and others.

Per rai expenditures of pesticide and herbicide are shown in table 4.7.

Table 4.7. Expenditure of Pesticide and Herbicide (baht/rai)

Crops	Mean	Std. Dev.	Coeff. Dev.	Minimum	Maximum
Pesticide Expenditure					
Rice	84.5	59.199	91.78	0	253.0
Soybean	123.0	48.267	39.25	0	243.0
Potato	481.2	245.24	53.18	85.0	1450.0
Tomato	376.1	122.90	32.68	208.0	618.0
Garlic	150.0	59.034	39.38	0	270.0
Herbicide Expenditure					
Rice	58.7	33.64	57.30	0	232.4
Soybean	114.7	59.59	51.95	0	187.5
Potato	80.1	103.05	128.65	0	400.0
Tomato	104.8	18.40	17.56	28.0	112.5
Garlic	101.5	13.87	13.67	42.0	110.0

Table 4.8. Number of farmers who do not apply chemical fertilizer, pesticide and herbicide

Crops	Rice	Soybean	Potato	Tomato	Garlic
Fertilizer	47	12	0	0	0
Pesticide	70	3	0	0	3
Herbicide	25	4	29	0	0

4.4.5. Seeds

Purchased and family stored seeds are regarded as having the same quality. Farmers tend to use the same

quantity of seeds per rai except for the potato crop. For rice, there are 111 farmers who report having used exactly 10 kg seeds per rai, 10 farmers report having used more than 10 kg per rai, and another 72 farmers use less than 10 kg per rai. On average, farmers use 8.8 kg of rice seeds per rai. Among the 70 observations of soybean growers, 60 use exactly 10 kg of seeds per rai, 7 use less than 10 kg, and 3 use more than 10 kg per rai. Averaged over the whole sample, farmers use 10.1 kg of soybean seeds per rai. All observed tomato and garlic farmers use the same amount of seeds per rai, 10 grams and 120 kg per rai, respectively. Since for these crops, seed is a minor input compared to other inputs, it is not significant when enter the model. Thus, it is omitted.

Potato production use potato tuber as seed. The tubers contribute about 30 percent of the total production cost. In Thailand, seed potato is imported from the Netherlands and for the purpose of controlling potato acreage, seed potato is sold to farmers on a quota basis. Wide variations of seed rate is observed among the sample farmers. The lowest rate is only 15 kg per rai, which may not be adequate for establishing a reasonable plant population. The maximum rate is 150 kg per rai, and the average is 76.1 kg per rai. The variable is coded as SEED.

4.5. The Dependent Variable

The economic outputs (yield) of the crop production are used as the dependent variable. Wide variations of rice yield are observed. On average, rice yield is 874.9 kg per rai, with standard deviation of 84.40 kg per rai. The lowest yield is 400 kg per rai, and the highest is 1028 kg per rai. Average soybean yield in the sample is 238.9 kg

per rai, which ranges from 189 to 290 kg per rai. Potato is the most varied crop in terms of input use and output produced. Average yield is 3230 kg per rai, which ranges from 1000 to 7000 kg per rai. Average tomato yield is 5,337 kg with minimum 4,575 kg and maximum 6,450 kg. Average garlic yield is 2530.0 kg per rai. Lowest and highest are 2130 and 2750 kg per rai, respectively.

In all models, the dependent variables are measured as kg of output per farm and coded as Q.

Table 4.10. Descriptive Statistics of Yield of Five Crops (kg/rai)

Crops	Mean	Std. Dev.	Coeff. Dev.	Minimum	Maximum
Rice	674.9	84.4	12.51	400.0	1028.0
Soybean	236.8	28.7	12.11	189.0	290.0
Potato	3230.8	356.0	11.02	1000.0	7000.0
Tomato	5337.3	397.1	7.44	4575.0	6450.0
Garlic	2530.0	199.0	7.87	2130.0	2750.0

4.6. The Model and Specification

The duality theory of production and cost enables us to look at the problem of cost efficiency in three ways, the cost function approach, the profit function approach, and the direct production function approach. Given any one of these three functions, one can get the other two functions by Shepherd's lemma. Details of derivation can be found in Silberberg (1978).

If the production is

$$Q = F(X_1, X_2, \dots, X_m) \dots\dots\dots 4.1$$

where Q is output and X_i ($i=1,2, \dots, m$) is a vector of inputs, then the cost and profit functions are,

$$C = G(Q, r_1, r_2, \dots, r_m) \dots\dots\dots 4.2$$

and, $\Pi = H(p, r_1, r_2, \dots, r_m) \dots\dots\dots 4.3$

respectively. Where C is the total cost, p is the output price, and r_i is a vector of input prices, and Π is the total profit.

In the empirical analysis, none of these three functions has any absolute advantage over the other ones. The basis for choosing any of these approaches depends mainly on data availability and the nature of the data set. If we use the conventional production approach, because input quantity and output level are both endogenous variables, that is, they are decision variables based on output and input prices, subject to production technology available, simultaneous equations should be employed. On the other hand, profit or cost are functions of input and output prices, which in the market economy, are exogenous to the firm, so, single equation model is appropriate.

However, if the study area is relatively small, and input prices do not vary, or do not vary much across firms, then, multicollinearity will be a serious problem. In this case, the conventional production function approach would be appropriate.

In this study, the study area is confined to Nong Han subdistrict, San Sai district of Chiang Mai Province, so

there is no substantial difference in such input prices as land (for tenants) fertilizer, pesticide and herbicide among farmers. So, the conventional production function is estimated to get the parameters of productions. Single equation is used because farmers face the same input and output prices.

4.6.1. Model Specification

In this study, the production functions are assumed to be Cobb-Douglas, which has been widely used in agricultural studies. For each crop, a group of homogeneous Cobb-Douglas production functions are specified and estimated by both OLS and MLE. Selection for the final models is based on the significant levels of the estimated coefficients and the estimability of frontier productions². If the specification gives more significant coefficients and at the same time, the corresponding frontier production function parameters are estimated, this model is retained for final analysis. The final estimation model for each crop is listed below.

Where ONE is the constant term; Q, LD, LB, PEST and HERB are total output (kg per farm), land size (planted area, rai), labor input in man-days, pesticide and herbicide expenditure (baht), respectively, all measured on a per farm basis.

2. As discussed in the previous chapter, some model specification may result in the error being dominated by the random component, and OLS and maximum likelihood estimations yielding the same estimates.

Table 4.11. Model Specification

Crops	Model Specification
Rice	$\ln Q = (\text{ONE} + \ln LD + \ln LB + \text{LN FERT1} + \ln \text{FERT2} + \ln \text{PEST} + \ln \text{HERB}) \beta_1 + (v-u)$
Soybean	$\ln Q = (\text{ONE} + \ln LD + \ln LB + \text{LN FERT1} + \ln \text{FERT2} + \ln \text{PEST} + \ln \text{HERB}) \beta_1 + (v-u)$
Potato	$\ln Q = (\text{ONE} + \ln LD + \ln LB + \ln \text{FERT1} + \ln \text{PEST} + \ln \text{HERB} + \ln \text{SEED}) \beta_1 + (v-u)$
Tomato	$\ln Q = (\text{ONE} + \ln LD + \ln LB + \ln \text{FERT1} + \ln \text{FERT2} + \ln \text{FERT3} + \ln \text{PEST} + \ln \text{HERB}) \beta_1 + (v-u)$
Garlic	$\ln Q = (\text{ONE} + \ln LD + \ln LB + \ln \text{FERT1} + \ln \text{FERT2} + \ln \text{PEST} + \ln \text{HERB}) \beta_1 + (v-u)$

FERT1 is the pure N which comes from urea and 21-0-0 for the rice model; 11-8-6 for the soybean model; 12-24-12 for the potato model, 13-13-21 for the tomato model; the sum of effective elements of N and P for the garlic model. All are measured in kg per farm.

FERT2 is also measured in kg per farm, which is fertilizer 16-20-0 for rice; 12-24-12 for soybean; 15-15-15 for tomato and potassium (K_2O) for garlic. It does not appear in the potato crop, since farmers only use one formula of fertilizer.

FERT3 is urea applied in tomato production, kg per farm.

4.6.2. Frontier Production Function Estimation

For a given sample, the frontier production function reflects the maximum products producible at a given level of inputs as defined by the sample's "best practice" farms. In this study, farms are assumed to maximize outputs at the given level of inputs.

The model is written in its general form as,

$$\ln Q = f(\ln X, \beta) + \epsilon, \quad \epsilon = v - u \quad \dots\dots\dots 4.4$$

where, Q is output of each crop per farm in kg,

X is a vector of input as listed in the above section,

β is a vector of the parameters to be estimated,

v is a two-sided error term with zero mean and normal distribution.

u ($u \geq 0$) is the one-sided error term to capture the "technical inefficiency" of the firms.

The production function is first estimated by OLS method. Breusch and Pagan's method is used to test heteroscedasticity. Chi-square test shows that there exists heteroscedasticity in rice and tomato models at 5 percent confidence level, therefore, the Generalized Least Squares (GLS) method of correcting heteroscedasticity as proposed by White (1978) is employed to get the production parameters. Tests of returns to scale are conducted. Then, frontier production functions of the same specification are estimated via the Maximum Likelihood Estimation. All the processes are essentially done in LIMDEP³.

³ LIMDEP is an econometric program written by W.H. Greene. Copyright by W.H. Greene, Cornell University, NY 14853, USA.

4.6.3. Calculating Technical Efficiency and Identifying Factors Associated with Technical Efficiency

Technical inefficiencies of the individual farms and the sample mean inefficiency are calculated using Eqs. 3.6 and 3.5, and then translated into technical efficiency indexes. For a Cobb-Douglas production function, the technical efficiency index (TE) is calculated as,

$$TE_j = \exp(-u_j) \dots\dots\dots 4.5$$

The final step of the analysis involves identifying factors determining technical efficiency. This analysis is conducted only for the crops of which frontier production functions are estimated, namely, rice, tomato and garlic. A logit model is employed. Following Pindyck and Rubinfeld (1981), the model is written as,

$$\log \frac{TE_j}{1 - TE_j} = Z_j = \alpha + \beta X_j \dots\dots\dots 4.6.$$

where TE_j is the technical efficiency as calculated through 4.5, X is a vector of variables as defined in table 4.12, and coded in table 4.13, and the subscript j denotes j^{th} observation. Since the data on technical efficiency is continuous, as suggested by Pindyck and Rubinfeld (1981), OLS method is used to estimate the parameters of eq. 4.6.

Copyright © by Chiang Mai University
All rights reserved

Table 4.12. Definition of X vector for equation 4.8.

Rice:	X = ONE, LD, LB, DIST, POTATO, TOMATO, GARLIC, TENANT
Tomato:	X = ONE, LD, LB, DIST, FIXED, YEAR1
Garlic:	X = ONE, LD, LB, DIST, FIXED, YEAR2

Table 4.13. Coding of Determinants in the Logit Model

	1	0
POTATO	Potato rice cropping systems	Otherwise
TOMATO	Tomato rice cropping systems	Otherwise
GARLIC	Garlic rice cropping systems	Otherwise
TENANT	Tenant farmers	Owner-operator
LD	Operation size of the crop concerned in rai	
LB	Labor input in mandays	
DIST	Distance of farm field from farm family in km	
YEAR1	Years of growing tomato	
YEAR2	Years of growing garlic	
ONE	Constant term	

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
 Copyright © by Chiang Mai University
 All rights reserved