Information System impacts on Korean pig farm: productivity gain and its determinants

Ik Hoon Jang¹, Sung Hee Park², So Hyun Kim³ and Young Chan Choe⁴
Dept. of Agricultural Economics and Rural Development, Seoul National University
¹First author, iks0404@snu.ac.kr, ²sunghehiru@gmail.com, ³isolatedzany@snu.ac.kr, ⁴Corresponding author, agji@snu.ac.kr

Abstract

Korean pig farms face challenges from free trade environment. Many farms have adopted Information System to improve the productivity on the farm. This study applies statistical analysis to verify the productivity gain and its determinants by using information system on the pig farm. Using the farm level weekly data from the PigPlan from 2006 to 2012, we find productivity gain on the farm increases gradually after adoption of the IS on the farm. We also find time of IS usage, the size of farm, farrowing rate, and total piglet per litter link to the productivity gains.

Keywords: determinant factors, pig breeding management, agricultural information system, decision support, pigs per sow per year (PSY)

1. Introduction

Information system (IS) provides better decision making and systematic management for the agri-business. It also pays off to the farm by increasing productivity and profit. However, direct impact of information of the farm is difficult to measure. The measuring the impact of information system on the farm is not an easy task.

Electronic data recording and processing facilities have developed rapidly in the swine and dairy sectors[1]. Also, the number of farms adopting information system and information technology has been increased steadily [2,3]. However only a few studies have quantified the effect of IS [4-14].

This study intends to find impact of information system on agricultural industry of the pig farms, especially for productivity gain and its determinants in pig industry. We use multiple regressions and decision tree method on farm production data from PigPlan, the most widely used IS on Korean pig farm. This study applies time series panel data to estimate the productivity gains from 2006 to 2012.

This study applies time panel data analysis to estimate the productivity gains from 2006 to 2012. In the next section, environment for Korean pig sector is discussed. The third section reviews literature on IS impact of pig industry. The fourth section explains results from the analysis and is followed by final section to conclude the study.

2. Korean Pig Sector

The swine industry takes about 12% of farm by production in Korea [15]. The size of pig industry tripled in last 30 years. In 2014, there are 9,500,000 pigs on average of the farm and 750,000 tons of pork production (Korea hog raisers’ association). Even after Food and Mouth Disease (FMD), the swine industry still takes 11% of the total agricultural production and
30% of livestock production value. It produced 4.5 trillion won in 2012 and self-sufficiency rate reaches 79% [16].

Farms with 1,000 pigs or more have 88% of total production. Farms take about 20% of overall pig farms and contracted farms are expected to grow because of the swine industry cooperative which desires to secure competitiveness and government support.

Free Trade Agreement (FTA) between Korea and other countries including United States and European Union also influence Korean swine industry. Compared to swine industry in Korea, U.S. and EU have much larger scale and vertically integrated swine industry. The increase feed cost and the reinforcement of environmental regulation will discourage the productivity. The government also discourages the pork price to stabilize consumer prices.

Various Agricultural information systems were developed by agricultural agencies, feed grain companies, and private software developers and were diffused in the late 1990s. Although the introduction of MIS was relatively late in Korea, various information systems including enterprise resource planning (ERP) has been adopted.

PigPlan was developed by the Agricultural Information System Lab at Seoul National University in 1995. PigPlan is the most widely used program in Korea. For pig farms, PigPlan supports feeding management, disease control and the prevention of epidemics through the management of herd breeding and pig growing. The introduction of production management information system allows data to be analyzed for various purposes. Fig 1 shows steadily increased number of PigPlan users.

**Figure 1. Number of farms using PigPlan from 2000 to 2013**

The PigPlan at pig breeding farms manages performance of each sow pigs and help sales management, parceling-out of breeding pigs, boar semen, and business administration. PigPlan also supports for general farm management.

3. Literature Review

The effects of IS are regarded as an important aspect when it comes to decision making at the adoption phase on the possibility of recovering investment cost with extra revenue [6].

Electronic data recording and processing facilities have been developed rapidly in the swine and dairy sectors [1]. Although the use of information system has increased, not many studies have measured the impact of MIS. The direct impact from the adoption of information technology (IT) is difficult to measure. [7], show MIS benefits improved decision making.
[13], measured the impact of IS and tested the ‘productivity paradox’ of IS. They applied Restricted Maximum likelihood Estimation (RMLE) on data from 81 farms that adopted PigPlan system in Korea. They found a positive productivity improvement using IS on swine farm. PigPlan system increased 0.52 in PSY (pigs per sow per year) and decreased 0.087 in the number of litter per sow per year (LSY).

Table 1. Factors Listed in Literatures to Affect Swine Productivity by IS

<table>
<thead>
<tr>
<th>Authors</th>
<th>Method</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>[4]</td>
<td>Cross-Sectional Data, Correlation Analysis</td>
<td>Total litter size for sow per year (the number of unproductive days, average weaning age), weaning success rate (birch accident rate, mortality rate before weaning)</td>
</tr>
<tr>
<td>[5]</td>
<td>Cross-Sectional Data, Correlation Analysis</td>
<td>Total litter size for sow per year, the number of unproductive days, farrowing rate, mortality rate before weaning</td>
</tr>
<tr>
<td>[9]</td>
<td>Cross-Sectional Data, Correlation Analysis</td>
<td>the number of suckling days, candidate pig mortality rate, the rate of onset of estrus, sow mortality rate</td>
</tr>
<tr>
<td>[11]</td>
<td>Cross-Sectional Data, Correlation Analysis, T-test</td>
<td>Total litter size per year, PSY, the number of mortality heads, the number of suckling days, farrowing intervals</td>
</tr>
<tr>
<td>[10]</td>
<td>Cross-Sectional Data, Correlation Analysis, Regression Analysis</td>
<td>The number of weaned pigs(pigpen form, heat conservation facility, weaning age, sterilizing pigsty, sow rotating rate)</td>
</tr>
<tr>
<td>[12]</td>
<td>Cross-Sectional Data, X2 test</td>
<td>The number of weaned pigs(weekly management, 2-3site pigpen management, species unification, liquid feed, artificial fertilization)</td>
</tr>
</tbody>
</table>

[14] analyzed the production process of pig breeding and the key points affecting the pork quality. The study adopted radio-frequency identification (RFID) technology to identify retrospective coding, and established many levels of pig breeding management systems for the whole quality tracing and breeding process monitoring. After the system was implemented, personnel management fees decreased by 5%, feed costs decreased by 8%, and the eligibility ratio of pig inspections increased by 10%.

Table 1 shows the previous studies that identify the significant factors affecting pig production. Only few studies were conducted to find the determinants. Furthermore, they used the cross sectional data and time series data. In this study, we use the time series panel data for the first time to find IS impact on pig production.

4. Date and Methods

4.1. Data

We collected data from swine farms who have adopted information systems (PigPlan) which is most widely employed. Approximately 35.8% of swine in Korea has been enrolled in this system. The total number of observations is 1,366 on yearly farm level.

We have applied panel data methods, a yearly farm level, to estimate the productivity gains after adopting PigPlan. By merging farm level data collected, we obtained a time series panel
data set with 1,366 observations from 251 farms over the period 2006-2012. Earlier studies used cross-sectional data. We use the time series panel data at first time.

Following model is constructed to estimate the effect of adopting IS on the dependent variable, which is the number of piglets raised per sow per year (PSY).

\[
PSY_{it} = \alpha + \beta_1 \text{Exp}_{it} + \beta_2 \text{Scale}_{it} + \beta_3 \text{Del_rate}_{it} + \beta_4 \text{Total_birth}_{it} + \beta_5 \text{Day_re}_{it} + u_i + e_{it}
\]

\[i = 1, 2, \ldots, 251 \text{ (farm)} \quad t = 1, 2, \ldots, 7 \text{(time of IS usage)}
\]

Table 2 five dependent variables: time of IS usage, the size of farm, delivery rate, total piglet per litter, and percent of 7 days return of estrus. These variables were coded and employed in the analysis. All statistical analyses were performed using STATA version 10.0.

**Table 2. Definition of Variables**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coding</th>
<th>Variables (unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>y</td>
<td>Pigs per sow per year : PSY</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>Use</td>
<td>Time of IS usage (year)</td>
</tr>
<tr>
<td></td>
<td>Scale</td>
<td>The size of farm (the number of sow)</td>
</tr>
<tr>
<td></td>
<td>Farrowing</td>
<td>Farrowing rate</td>
</tr>
<tr>
<td></td>
<td>LitterNo.</td>
<td>Total piglet per litter (no. of piglet)</td>
</tr>
<tr>
<td></td>
<td>Estrus</td>
<td>Percent of 7 day return of Estrus (%)</td>
</tr>
</tbody>
</table>

**4.2. Decision tree Analysis**

To find factors that affect swine productivity, this study uses decision tree, which is a type of data mining technique. Decision tree bases on data and automatically separates dependent variables into homogeneous groups. There are various algorithms invented, and they generally start from the root node which contains entire data and form a tree with branches and leaves in a systematic way. Each algorithm creates offspring branches from parent branch from split criterion. Offspring branches then become a parent branch and repeat the same process. This process continues until every node reaches to the end [17].

There are several algorithms that present the criterion that distinguishes target variable’s distribution such as by informing users which input variables and input variable value points need to be used.

CHAID(Chi-squared Automatic Interaction Detection) is used to find statistical relationship between two variables. Different from CART, CHAID stops forming tree before overfitting data. CHAID first transforms consecutive predictors into categorical predictors by grouping consecutive predictors for each section. Based on predictors’ and target variables’ contingency table, CHAID then finds chi-square statistics and selects the most significant variable and use it for a separation variable.

CART(Classification and Regression tree) algorithm first classifies groups based on specific variable and separates groups by calculating the probability of choosing variables in other groups from a classified group. As the segment is pure, Gini’s coefficient decreases which also reduce the probability. In this algorithm, variables are selected based on Gini’s coefficient.

C4.5 algorithm is based on entropy coefficient that measures the purity of nodes. The classification is based on variables that have the smallest entropy. CART forms duality division for each node and creates binary separation tree structure. C4.5 also forms binary
separation for consecutive predictors. However, C4.5 creates ramous separation structure for nominal predictors.

Similar to C4.5, QUEST(Quick Unbiased Efficient Statistical Tree) can only analyze nominal target variable. Division method conducts binary separation by using different separating criteria depending on the criterion of predictors. For separation variable selection, if predictors are consecutive, ANOVA, F-test or Levene test are used. If predictors are nominal, Pearson’s chi-squared test is used to select the smallest significant probability separation variable [18, 19].

A decision tree is sensitive to sample size and has less ability to analyze consecutive data compared to metric model such as regression analysis. However, a decision tree is a powerful tool in classification and prediction.

First, decision trees have more freedom from various statistical hypothesis compared to regression analysis and other metric models which use production predicting approach. Metric model is generally based on characteristic and data structure of input variable. On the other hand, decision trees are not based on the correlation between independent variables and distribution. Thus, they can increase the outcome of prediction model and have advantage in general model.

Second, decision trees can find variable that has nonlinearity effect. Regression analysis and econometric models assume linearity of data, so they can only find general effects by using standardized beta coefficient which indicates the most effective factors. However, decision tree can find effects that are based on certain conditions. For example, if decision tree branches out by certain scale, it can find other productivity improvement factors according to that scale.

Third, decision tree finds variables that are influenced by simple classification logic. Thus, it is easy to understand prediction basis and variables that influence group classification

5. Results

5.1. Multi Regression Analysis

<table>
<thead>
<tr>
<th></th>
<th>Time of IS usage</th>
<th>PSY</th>
<th>The size of farm</th>
<th>Total piglet per litter</th>
<th>Percent of 7 day return of Estrus</th>
<th>Farrowing rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1366</td>
<td>1366</td>
<td>1366</td>
<td>1366</td>
<td>1366</td>
<td>1366</td>
</tr>
<tr>
<td>Mean</td>
<td>3.39</td>
<td>21.984</td>
<td>425.240</td>
<td>11.472</td>
<td>87.684</td>
<td>82.101</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.048</td>
<td>0.0554</td>
<td>10.6682</td>
<td>0.0190</td>
<td>0.1909</td>
<td>0.1867</td>
</tr>
<tr>
<td>medium</td>
<td>3.00</td>
<td>22.050</td>
<td>289.550</td>
<td>11.500</td>
<td>88.700</td>
<td>82.900</td>
</tr>
<tr>
<td>mode</td>
<td>1</td>
<td>22.0</td>
<td>115.0</td>
<td>11.4^a</td>
<td>91.5</td>
<td>80.5</td>
</tr>
<tr>
<td>Standard error</td>
<td>1.776</td>
<td>2.0463</td>
<td>394.2901</td>
<td>0.7028</td>
<td>7.0563</td>
<td>6.9016</td>
</tr>
<tr>
<td>variation</td>
<td>3.153</td>
<td>4.187</td>
<td>155464.664</td>
<td>.494</td>
<td>49.791</td>
<td>47.632</td>
</tr>
<tr>
<td>range</td>
<td>6</td>
<td>13.8</td>
<td>2985.9</td>
<td>5.2</td>
<td>51.3</td>
<td>48.4</td>
</tr>
<tr>
<td>minimum</td>
<td>1</td>
<td>14.1</td>
<td>39.1</td>
<td>9.1</td>
<td>49.8</td>
<td>51.6</td>
</tr>
<tr>
<td>maximum</td>
<td>7</td>
<td>27.9</td>
<td>3025.0</td>
<td>14.3</td>
<td>101.1</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 3 shows the descriptive statics of the research model. Total observations are 1,366 from 251 farms data base from 2006 to 2012. Average time of IS use is 3.39 years.
Figure 2 shows the yearly trends of PSY and a representative indicator of swine productivity. The PSY for average farms and biennial management survey on pig farms (Korea Pork Producers Association). The number of pig farms with IS steadily increased from 142 farms in 2003 to 251 farms in 2012.

![Figure 2. PSY Trend between IS users’ and Korean National Average](image)

Figure 3 shows the yearly PSY trend of the pig farms with IS, depending on the scale of sow, the time of IS use varies from four to seven years. The numbers of pig farms for each different sow pig scale are 82 farms for under 200 sow pigs, 118 farms for 200~500 sow pigs, 35 farms for 500~1000 sow pigs, and 16 farms for more than 1000 sow pigs. Figure 3 shows that PSY for all farms were increase according to their time of IS usage. However, 200~500 sow pig farms were not in direct proportion.

![Figure 3. PSY Trend by Year after Adoption of IS and by Scale of Farms](image)

The productivity has improved by specification management that efficiently used IS. In particular, big-scale farms with more than 1000 sow pigs showed the most significant productivity improvement with the IS. Therefore, IS can be more effective in large-size farm management than in small-size farm management.

Table 4 shows the results of multiple regression analysis. The model is statistically significant as the F value is 0.01 level. The model explains the PSY relatively well with 0.5266 R2 and 0.5271 adjusted R2. Both values are slightly higher than 0.5 which suggests a robust explanatory power of this model.
Table 4. Results of Multiple Regression

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Err</th>
<th>p-Value</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time of IS usage</td>
<td>0.0490</td>
<td>-2.23</td>
<td>0.026</td>
<td>0.0424</td>
</tr>
<tr>
<td>The size of farm</td>
<td>0.0002</td>
<td>2.02</td>
<td>0.044</td>
<td>0.0386</td>
</tr>
<tr>
<td>Del_rate</td>
<td>0.1292</td>
<td>21.17</td>
<td>0.000</td>
<td>0.4360</td>
</tr>
<tr>
<td>Total piglet per litter</td>
<td>0.9823</td>
<td>16.37</td>
<td>0.000</td>
<td>0.3374</td>
</tr>
<tr>
<td>Percent of 7 day return of Estrus</td>
<td>0.0597</td>
<td>10.77</td>
<td>0.000</td>
<td>0.2058</td>
</tr>
</tbody>
</table>

The effect of delivery rate is the greatest among predictor variables (b=0.4387) followed by the total litter size (b=0.3374) and the percent of return of estrus in 7 days (b=0.202). The effects of period of information system use and the size of the farm are comparatively smaller than others.

5.2. Results of decision making tree analysis
Figure 4 presents the analysis results of decision making tree on the size of farm, the length of information system use which hold indirect effects on PSY. The most effective predictor for the productivity of a farm is the size of a farm. A higher productivity is associated with longer IS use regardless farm size. On the other hand, in large farms, the time of usage makes a big difference on the IS impact. While the Time of IS usage is longer than 1 year, then PSY increases 0.765 PSY per year. The farms with the time of IS usage under 1 year show no impact on PSY. Overall the larger the farm size and the longer they use IS, the more PSY increased. The farm with more than 432.1 sows and time of IS usage more than 3 years benefits greater productivity gains at 1.467 PSY impact from IS adoption.

5.3. The Effects of Influence Factors

Figure 5 presents the analysis results of decision making tree on the farrowing rate, Total piglet per litter and the percent of 7 day estrus which hold direct effects on PSY.

As Figure 5 shows that farrowing rates carry the greatest effect on PSY. The farms with farrowing rate more than 82.7%, total piglets per litter more than 12.7% and percent of 7 day return of estrus more than 87.6% benefits the greatest productivity among all pig farms gains at 25.7 PSY impact from IS adopt in. The farms with farrowing rate more than 82.7% but total piglet per litter under 10.2 and percent of 7 day return of estrus under 10.2% gain at 20.111 PSY impact. This PSY improvement is the smallest impact among farms with farrowing rate more than 82.7%

The farms with farrowing rate under 63% benefit the smallest among all pig farms 17.761 PSY impact. And farms with farrowing rate more than 73.2% and under 82.7%, total piglets per litter more than 11.2 and percent of 7 day return of estrus more than 84.8% benefit 21.999 PSY impacts. This PSY improvement is the lowest impact among farms with farrowing rate under 82.7% The farms with farrowing over 82.7%, total piglet per litter more than 10.2 and percent of 7 day return more than 80.3% under 93% gains 21.904 PSY impacts.

In summary, Farms with farrowing rate more than 82.7%, total piglets per litter more than 12.7% and percent of 7 day return of estrus more than 87.6% presented the highest PSY improvement while with farrowing rate under 63% presented the lowest PSY improvement.

Note: y = PSY and n = the number of observations

Figure 5. Decision Tree by Influence Factors
6. Conclusion and Limitations

This study analyzed the IS (PigPlan) impact on pig farm. For the first time, this study used time series panel data set that consisted of 1,366 observations from 251 farms over the period of 2006-2012. Results of both multi regression and decision tree analysis show a yearly increase in the productivity of pig farms. Furthermore, big-scale farms with more than 1000 sow pigs showed the best improvement. Therefore, the application of IS program can be more effective in large-scale farm than small ones.

IT payoff is still a controversial issue among researchers and practitioners but this study verified the productivity gains after adopting IS. Recorded data on information system can be utilized widely and help improve systematic management in agriculture.

This study has a few limitations. The data do not include demographic information of users, which may have an impact. According to PSY trend of pig farm scale in Korea, productivity has been decreased from 2003 to 2009, and increased since then. On the other hand, the PSY trend of IS users has no big difference. Thus in order to clarify the productivity gains using IS, further work about the reason of that PSY difference between users and nonusers needs to be done and comparison between before and after adoption should be conducted.

Acknowledgments

This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the CITRC(Convergence Information Technology Research Center) support program (NIPA-2014-H0401-14-1008) supervised by the NIPA(National IT Industry Promotion Agency)

References


Authors

Ik- Hoon Jang, he received the BS in Department of Electrical and Electronic Engineering from Yonsei University in 2006. He received the MS in Department of Agricultural Economics and Rural Development from Seoul National University in 2012. He is currently a Ph.D course student in the Graduate school of Agricultural Economics and Rural Development, Seoul National University. His current research interests include agricultural information systems, e-business, and data mining.

Sung- Hee Park, she received her MS degrees from Seoul National University in Korea. Currently, she is a candidate for the Ph.D at the Agricultural Economics and Rural Development at the Seoul National University. Her research interests include information management system, environmental management, and food network.

So- Hyun Kim, she received the BS in Department of Agricultural Economics and Rural Development from Seoul National University in 2014. She is a candidate for the MS at the in the Graduate school of Agricultural Economics and Rural Development, Seoul National University. Her current research interests include agricultural information systems, econometrics for big data analysis.
Young-Chan Choe, he received the MS and Ph.D in Department of Agricultural Economics from Michigan State University in 1989 and 1991, respectively. He is currently a Professor in Seoul National University. His current research interests include agricultural information systems, e-business in the food and agricultural sector, and big data analysis.