The Impact of Palay Harvest Area, Inorganic Fertilizer Use, and Production Volume on Regional Rice Retail Prices Across Multiple Years: A Linear Regression and Correlation Analysis

Paul Matthew Gayatiña
paul_gayatina@dlsu.edu.ph
De La Salle University

Justin Christian Vilbar
justin_vilbar@dlsu.edu.ph
De La Salle University

Julia James Berin
julia_berin@dlsu.edu.ph
De La Salle University

Abstract: The study focuses on the factors influencing the costs of well-milled, regular-milled, and special-milled rice in the Philippines. Data from 2020 to 2023 were gathered from the Philippine Statistics Authority using multiple linear regression, with an emphasis on the Cordillera Administrative Region (CAR) and Regions I through V. The study examines the effects of regional and temporal changes in rice prices, as well as area harvested, production volume, and fertilizer application. After performing tests, results show that each rice type responds differently to production and market conditions. The well-milled rice model presented a negative correlation with production volume, indicating that high supply generally results in lower prices. Results also show that Regions II and III and yearly trends, particularly years 2020 and 2023, play a vital role in pricing. The regular-milled rice model, on the other hand, demonstrated the strongest explanatory power, with the variable harvested area negatively affecting prices, while Regions III and V and years 2022 and 2023 showed an increase in price per unit increase of their respective variables. Lastly, the special-milled rice model identified the fertilizer application as an important price determinant, with Region V, CAR, and the year 2020 showing a price decrease, compared to 2023's price increase. These results show the importance of regional and temporal factors in rice pricing, improving previous research on production-price relationships. The data provide valuable insights for the Department of Agriculture (DA) in designing effective price stabilization strategies. In line with this, farmers can also benefit from the study by using these results to optimize fertilizer application and possibly make a change in production strategies. Consumers will gain awareness of the price trends of rice, and future research could further improve these models by incorporating other variables such as weather patterns and economic factors.

Key Words: rice market analysis; rice pricing; agricultural economics; Philippine rice production; multiple linear regression; regression analysis; statistical modeling

1. INTRODUCTION

In Filipino culture, rice is treated as a necessity in each meal, and is considered a staple food in the Philippines. The country has been considered one of the top rice producers in the world, with the Department of Agriculture (2022) stating that the Philippines ranked 8th behind China, Indonesia, India, and other Southeast Asian nations. Research by institutions and colleges has found multiple factors that affect rice productivity and how these factors affect the rice price in our market. One of the factors is the total area harvested. Gomez (2024) states that Quarter 3 of 2024 has significantly decreased, dropping from 926,923.12 hectares (HA) in Q3 2023 to 792,221.21 HA, with a decrease of 134,701.91 HA. Furthermore, the study also indicates that the harvested area decreased by 11% between Q2 and Q3 of 2024, which led to a 13% drop in production. However, due to the rise in imports, rice prices have also dropped throughout the period. Furthermore, the decreased usage of subsidized fertilizers affects rice yield and the amount of rice harvested in Indonesia, according to Fahmid et al. (2022). Antonio (2024) stated that from 1994-2023, the rice price inflation varied across different regions of the Philippines. Regions that were insufficient in the rice sufficiency index were significantly higher than their counterparts. The insufficient regions were NCR, CALABARZON, Central Visayas, Zamboanga Peninsula, Northern Mindanao, and Davao Region, while other regions had sufficient rice.

The study is beneficial for government agencies like the DA to formulate effective rice price stabilization and food security planning policies or programs. The results will contribute in predicting future price inflation, and customers will find it helpful to understand how the production factors correlate to the price of rice. The regression models can be used to predict how changes in harvest area or fertilizer use impact consumer rice prices. Further, farmers can use the results to make more efficient decisions about fertilizer use on their crops.

Fitrawaty et al. (2023) state that Indonesia will have an inequality problem in the future because of its low rice production and high rice prices. The researchers ran a simulation on the production variable of domestic rice by increasing and decreasing the production level of domestic rice by ± 3.5% per year from the previous year. Additionally, one of their findings indicates that an increase in the volume of rice production will lower the rice price in the long run, and a decrease in the volume of rice production will increase the rice price in the long run. Furthermore, Beltran (2024) indicates that one of the factors that affects the farmgate price of rice is production growth. It also discussed how the farmgate price affects the wholesale and retail price of rice.

According to Qodri et al. (2023), the harvested area in Indonesia has a coefficient value of -1.12. This implies that there will be a -1.12% price increase in the following period for every 1% drop in the harvested area. Furthermore, research suggests that the price of rice and the area harvested are indirectly related. This finding implies that the price of rice increases when the harvested area decreases.

Rice is one of the most important agronomic crops of the Philippines. With these, fertilizer usage increased throughout the years. With 2.52 megametric tonnes, inorganic fertilizers were widely used in our nation to boost corn and rice yields (PSA 2024). Galang et al. (2024) state that our country has a problem with farmers' vast use of inorganic fertilizers since they perceive that high usage of fertilizers will lead to higher crop yields. Since fertilizers have an impact on both productivity and farmers' expenses, their use has been essential in rice farming. Similar to the Philippines, one of the nations with the highest rice production is Indonesia. Fahmid et al.(2022) indicate that the increase in the price of Indonesian rupiah-subsidized fertilizers will affect production. In a simulation of 300/kg, the national rice harvest area would reduce by 186,219 ha, the national rice productivity would drop by 0.09 tons/ha, and the national rice production would drop by 0.94 million tonnes of milled dry rice (MDR/GKG). These would affect production, which will eventually lead to an increase in price in the market.

The study is only limited to Luzon areas, specifically Regions I-V and the Cordillera Administrative Region (CAR). Luzon is the largest island group in the Philippines and also has important locations for the production of rice for the population. The analysis offers a four-year perspective on the correlations between factors and covers data from 2020 to 2023. Furthermore, the study only analyzes the variables, harvest area, production volume, and fertilizer use and their relationship with rice retail prices, eliminating other potential factors that might influence rice pricing. The study employs linear correlation and regression analysis to determine relationships between variables. The study only relies on the 2020-2023 data of the variables, which might not adequately reflect cyclical patterns or longer-term trends in the rice market.

This study aimed to determine the linear correlation and regression relationship between key agricultural factors and rice retail prices in the Philippines. Specifically, this research sought to answer the following questions:

1. What was the overall relationship between palay harvest area (in hectares), volume of rice production (in metric tons), area of fertilizer applied (in hectares), year(year of the data), region(region of the data) and with price of well-milled rice, regular-milled rice, and special rice per kilo?
2. Could a multiple linear regression model using these five factors effectively predict rice retail prices in the Philippines?
3. Among the five factors (harvest area, production volume, fertilizer use, year, or region), which were significant predictors of well-milled rice, regular-milled rice, and special rice prices?
4. How much of the total variation in rice prices could be explained by the regression model?

2. METHODOLOGY

The study utilized databases acquired from the Philippine Statistics Authority. The specific databases used were *Cereals: Retail Prices of Agricultural Commodities by Geolocation, Commodity, Year and Period*, *Palay and Corn: Volume of Production in Metric Tons by Ecosystem/Croptype, by Quarter, by Semester, by Region, 1987-2024*, *Palay and Corn: Area Harvested in Hectares by Ecosystem/Croptype, by Quarter, by Semester, by Region and by Province, 1987-2024*, and *Palay: Estimated Inorganic Fertilizer Use by Geolocation, Grade and Year by Region, Area Harvested and Year by PHILIPPINES, Area Harvested and Year by PHILIPPINES, FERTILIZER and Year*. The researchers extracted data from 2020 to 2023, focusing on Regions I to V and the Cordillera Administrative Region (CAR).

The variables that were used in the study are the following:

Table 1. Variables Used in the Study

Variable	Data Type	Description
wellMilled	Float	Price in Philippine peso of well-milled rice per kilo
regularMilled	Float	Price in Philippine peso of regular-milled rice per kilo
specialMilled	Float	Price in Philippine peso of special-milled rice per kilo
volumePalay	Float	Volume of rice production in metric tons
areaPalay	Float	Area of palay harvested in Hectares
areaApplied	Float	Area of fertilizer applied in Hectares
region	Nominal	Region where rice was harvested 1 if Region is Region I (Ilocos Region) 2 if Region is Region II (Cagayan Valley) 3 if Region is Region III (Central Luzon) 4 if Region is Region IV-A (Calabarzon) 5 if Region is Region V (Bicol) 6 if Region is Cordillera Administrative Region
region2	Boolean	Dummy Variable for region 1 if the region is Region II (Cagayan Valley 0 otherwise
region3	Boolean	Dummy Variable for region 1 if the region is Region III (Central Luzon) 0 otherwise
region4	Boolean	Dummy Variable for region 1 if the region is Region IV-A (Calabarzon) 0 otherwise
region5	Boolean	Dummy Variable for region 1 if the region is Region V

		(Bicol) 0 otherwise
region6	Boolean	Dummy Variable for region 1 if the region is Cordillera Administrative Region 0 otherwise
year	Nominal	Year when rice was harvested 1 if Year is 2021 2 if Year is 2022 3 if Year is 2023 4 if Year is 2020
year2	Boolean	Dummy Variable for year 1 if the year is 2022 0 otherwise
year3	Boolean	Dummy Variable for year 1 if the year is 2023 0 otherwise
year4	Boolean	Dummy variable for year 1 if the year is 2020 0 otherwise

The categorical variable region had 6 categories to indicate which region the rice was harvested from. However, to make this variable processable for the multiple linear regression model, it was converted into 5 dummy variables with region as the base category. The categorical variable year was also a nominal variable with 4 categories to signify the year the rice was harvested. This was also converted into 3 dummy variables with year as the base category. These base category variables were then explicitly not included in the multiple linear regression model.

To determine the relationships of the independent predictor variables with the predictor variable, the correlation coefficients are computed and interpreted to determine the degree of their relationship and whether this relationship is significant or not.

For further understanding of the relationship of these variables, a multivariate analysis was used. The regression model will determine the predicted value of rice retail prices depending on the area harvested in hectares, volume of production in metric tons, estimated inorganic fertilizer use, the region where rice is harvested, and the year when rice is harvested.

The regression model was modelled to be

$$\widehat{Y}_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{11} X_{11} + \epsilon \quad \text{(Eq. 1)}$$

Where

$\widehat{Y}_i$ is the predicted value of the i^{th} rice retail price

β_0 is the y-intercept

β_1 is the slope associated with the area harvested variable

X_1 is the area harvested variable

β_2 is the slope associated with the volume of production variable

X_2 is the volume of production variable

β_3 is the slope associated with the estimated inorganic fertilizer use variable

X_3 is the estimated inorganic fertilizer use variable

β_4 is the slope associated with the first region dummy variable

X_4 is the dummy variable associated with region II

β_5 is the slope associated with the second region dummy variable

X_5 is the dummy variable associated with region III

β_6 is the slope associated with the third region dummy variable

X_6 is the dummy variable associated with region IV-A

β_7 is the slope associated with the fourth region dummy variable

X_7 is the dummy variable associated with the region V

β_8 is the slope associated with the fifth region dummy variable

X_8 is the dummy variable associated with Cordillera Administrative Region

β_9 is the slope associated with the first year dummy variable

X_9 is the dummy variable associated with the year 2022

β_{10} is the slope associated with the second year dummy variable

X_{10} is the dummy variable associated with the year 2023

β_{11} is the slope associated with the third year dummy variable

X_{11} is the dummy variable associated with the year 2020

ϵ is the random error component

To decide on what model to use for each rice type, the variables were subjected to automatic search procedures, specifically the forward selection, backward elimination, and stepwise selection procedures. The four models, including the full-rank model, are then compared with one another.

The F-test was used to compare the significance of the models. The characteristics of the models were also observed, particularly the Mean Squared Error (MSE), Mallows' C_p (C(p)), Akaike's Information Criterion (AIC), Schwarz Bayesian Criterion (SBC), and Prediction Sum of Squares Criterion (PRESS) values. The models were also checked to see if they had multicollinear variables.

After comparing all the factors of the models with one another, a singular model is taken for each of the three rice types and is interpreted for the study.

After which, tests for the significance of the predictors were used to determine which of the predictor variables are significant and insignificant. The variation in the dependent variable rice retail prices was also assessed by observing the adjusted R^2. The assumptions of a multiple linear regression model were also tested, specifically, the test for normality, the test for homoscedasticity, checking whether the mean of the probability distribution of errors is 0, and the test for independence.

All results were generated using the Statistical Analysis System (SAS) software.

3. RESULTS AND DISCUSSION

3.1 WELL-MILLED RICE

Well-milled rice prices exhibit moderate positive correlations with variables volumePalay, areaPalay, and areaApplied with pearson correlation coefficients of 0.306 and 0.292, and 0.295, respectively, suggesting a slight increase in rice prices as these agricultural production metrics rise. However, these correlations are not statistically significant, since their p-values are between 0.146 and 0.166.

Well-milled rice prices were modeled with 11 predictor variables. The full rank model resulted in a PRESS value of 14.09886, an MSE value of 0.27393, C(p) of 12.0000 with p=11, AIC of -23.7130, and SBC of -9.57634. The adjusted R^2 value was calculated to be 0.9421, and the model was shown to have multicollinear variables. However, the forward selection, backward elimination, and stepwise selection models all resulted in the same model. This model had a PRESS of 0.934055, MSE of 0.27973, C(p) value of 6.3814 with p=5, AIC of -25.4786, and SBC value of -18.41026. The model had an adjusted R^2 of 0.9409 and was determined not to be multicollinear. It can be observed that the MSE and C(p) values of the full-rank model are better characteristics as the MSE is smaller and the C(p) is closer to its p, however, the automatic selection models have lower values for PRESS, AIC, and SBC. The full-rank model is also seen to have a higher adjusted R^2, but the automatic selection model is Multicollinear. The researchers then determined the automatic selection model to be a better multiple linear regression model for the study.

The final multiple linear model for well-milled rice prices is

$$\widehat{wellMilled} = 43.33957 - 0.00000114(volumePalay) + 4.86801(region2) + 4.88181(region3) + 3.60246(year3) - 0.61373(year4) \quad \text{(Eq. 2)}$$

After going through the automatic selection process, the predictor variables left in the model are the volume of production, the dummy variable associated with region II, the dummy variable associated with region III, the dummy variable associated with the year 2023, and the dummy variable associated with the year 2020.

The model shows that with regard to the variable volumePalay, the price of well-milled rice is expected to decrease by 0.00000114 units for every unit increase in volume of rice production. The price of well-milled rice is also expected to decrease when the area is harvested in the year 2020; the expected decrease is 0.61373 units. If the rice is harvested in the year 2023 instead, the expected increase in well-milled rice price is 3.60246 units. If the rice is harvested from Region II, Cagayan Valley, the price of well-milled rice is expected to increase by 4.86801 units. However, when rice is harvested from Region III, Central Luzon, well-milled rice is expected to increase by 4.88181 units. All expected increased and decreased values only hold true if all other predictor variables are held constant.

The f-statistic to test for the significance of this model was calculated to be 74.25 with a p-value of less than 0.0001. At a 0.05 level of significance, it can be concluded that the model is a good fit for the data, thus significant. For the predictors, all five predictor variables had p-values less than 0.0001, thus are all significant at the 0.05 level of significance. The model had an adjusted R^2 value of 0.9409, which shows that 94.09% of the total variation in Well-Milled Rice prices can be explained by the volume of rice production, whether the rice was from region 2, whether the price was from region 3, whether the price year 3 year 4, after adjusting for the number of explanatory variables and sample size 24.

The model was checked as to whether it satisfied the assumptions of a multiple linear regression model. For the test of normality, the Shapiro-Wilk test statistic value was 0.957348 with a p-value of 0.3875, which signifies that the errors are normally distributed. As for the mean of the probability distribution of the residuals, the expected value of the errors was computed to be 0. The results for the test for homoscedasticity showed that the test statistic value was 2.00 with a p-value of 0.1458, which indicates that the model is homoscedastic. The test for independence resulted in the test statistic value of 2.356, with p-values for Pr<DW and Pr>DW being 0.7240 and 0.2760, respectively, thus indicating that the errors are independent. Thus, all assumptions of a multiple linear regression model were satisfied.

3.2 REGULAR-MILLED RICE

Prices for regular-milled rice display a moderate negative correlation, with Pearson correlation coefficients valuing around -0.30 to -0.33 for variables volumePalay, areaPalay, and areaApplied. This suggests that increases in these agricultural factors are linked to slight decreases in rice prices. However, these connections are not statistically significant, with p-values ranging from 0.11 to 0.15.

Regular-milled rice prices were examined using 11 predictor variables across various regression modeling techniques, including full-rank, forward/stepwise selection, and backward elimination. The backward elimination model emerged as the most effective in meeting the required assumptions, despite initially displaying issues with homoscedasticity. To address this, transformations were applied by raising the specialMilled variable to the power of three. These modifications allowed both the forward/stepwise and backward elimination models to satisfy all necessary assumptions.

When comparing model selection criteria with the new transformed data, the full-rank model's C(p) value of 12 was close to the p of 11, indicating a reasonable fit. However, the backward elimination model was ultimately chosen because it satisfied all the necessary assumptions, making it a more reliable option despite the full-rank model's acceptable C(p) value, with a C(p) value of 8.7237 with p=5. Also, the backward elimination model has a high adjusted R^2 of 0.9737, indicating that the model is a good fit. Furthermore, all of the assumptions, such as having small PRESS, AIC, and SBC, are satisfied.

The resulting multiple linear regression model is

$$\widehat{regularMilled^3} = 58113 - 0.02533(areaPalay) + 12976(region\,3) + 6668(region\,5) + 2744.50\, year\,2 + 19736(year3) \quad \text{(Eq. 3)}$$

The MLRM above shows that there is an increase for every unit decrease of areaPalay (area of Palay harvested), a increase for every unit increase of region3 (region III), an increase for every unit increase of region5 (region V), an increase for every unit increase of year2 (year 2022), and a decrease for every unit increase of year3 (year 2023).

Following the backward elimination process, the retained variables included areaPalay (area of Palay harvested), region3 (Region III), region5 (V), year2 (representing the year 2022), and year3 (representing the year 2023).

The model achieved an adjusted R^2 value of 0.9737, indicating that 97.37% of the total variation in regular-milled rice prices is accounted for by the variables areaPalay (the area of Palay harvested), region3 (Region III), region5 (Region V), year 2 (2022), and year 3 (2023), after considering the number of explanatory variables and a sample size of 23.

Following the significance tests, all variables were found to be significant predictors of the regular-milled rice price. Additionally, tests assessing the model's suitability for the data confirmed that it met all assumptions: the residuals were normally distributed, their mean equaled zero, they exhibited homoscedasticity, and they were independent.

3.3 SPECIAL-MILLED RICE

The correlation coefficients between the price of special-milled rice and the volume of palay production, area of palay harvested, and area applied of fertilizers show a positive correlation with the price of special-milled rice, with correlation coefficients above 0.5 for every variable. The p-value of these variables is observed to be less than 0.05, indicating that all variables are statistically significant.

To further investigate price determinants, special-milled rice prices were analyzed using 11 predictor variables across multiple regression modeling approaches: full-rank, forward/stepwise selection, and backward elimination. Among these, the backward elimination model best satisfied the necessary assumptions, though it initially exhibited multicollinearity. To resolve this, transformations were applied—specifically, converting the specialMilled variable to its natural logarithm form and squaring the areaApplied variable. These adjustments resulted in both the forward/stepwise and backward elimination models meeting all assumptions.

Comparing model selection criteria, the backward elimination approach demonstrated better alignment with statistical benchmarks. Its C(p) value of 7.9838 (compared to p = 5) showed a closer fit than the forward/stepwise model's C(p) of 11.1248 (p = 4), leading to the final decision favoring the backward elimination model.

The resulting multiple linear regression model is:

$$\begin{aligned} ln(\widehat{specialMilled}) &= 3.92427 + (1.16602E - 13) \\ &+ (areaApplied)^2 - 0.04667(region5) \\ &- 0.03885(region6) + 0.06625(year3) \\ &- 0.02442(year4) \end{aligned} \quad \text{(Eq. 4)}$$

The MLRM above shows that there is an increase for every unit increase of areaApplied (area applied of fertilizers), a decrease for every unit increase of region5 (region V), a decrease for every unit increase of region6 (CAR), an increase for every unit increase of year3 (year 2023), and a decrease for every unit increase of year4 (year 2020).

After performing backward elimination, the variables that remained were areaApplied2 (squared area applied of fertilizers), region5 (region V), region6 (CAR), year3 (year 2023), and year4 (year 2020).

The model had an adjusted R2 value of 0.8755, which shows that 87.42% of the total variation in Special-Milled Rice prices can be explained by areaApplied 2 (squared area applied of fertilizers), region5 (region V), region6 (CAR), year3 (year 2023), and year4 (year 2020), after adjusting for the number of explanatory variables and sample size 23.

After performing the significance test for all variables, it was observed that all variables are significant predictors of the price of special-milled rice. Furthermore, tests to examine whether the model was appropriate for the data were also performed. It is found that the model satisfied all assumptions: residuals follow a normal distribution, their mean sums to zero, they show homoscedasticity, and they are independent.

3. CONCLUSION

This research examined the determinants of rice prices for three quality grades—well-milled, regular-milled, and special-milled rice—using multiple linear regression methods to find significant predictors. The results indicated varying patterns of influence by rice type, and some important findings from the models.

For well-milled rice, the model produced an adjusted R^2 value of 0.9409, showing that 94.09% of price fluctuation is explained by volume of production and regional/temporal influences. More specifically, the negative volumePalay coefficient implies higher volumes of production reduce prices of well-milled rice slightly, whereas Regions II and III and years 2020 and 2023 have significant influences on prices. This model was validated for data analysis by meeting all statistical assumptions.

The regular-milled rice model performed even more explanatory power with the adjusted R^2 of 0.9737. Upon transformation to meet statistical demands, we obtained the result that areaPalay had an inverse relationship with regular-milled rice prices. There were strong regional effects, with Region III and Region V each registering strong positive price effects. Years 2022 and 2023 also had a high influence on the pricing behavior, implying significant temporal trends in this market segment.

For special-milled rice, the log-transformed model had an adjusted R^2 of 0.8755 with the squared area of fertilizer application ($areaApplied^2$) as a positive significant predictor. Interestingly, regional effects were mixed, with Regions V and CAR (Region VI) having negative coefficients, while temporal factors indicated variation, with 2023 having positive and 2020 having negative effects on prices.

In all three categories of rice, regional and temporal considerations were repeatedly found to be strong determinants, implying that geographical conditions and year-specific factors are important influences on rice prices in the Philippines. The relative importance of production indicators (volumePalay for well-milled, areaPalay for regular-milled, and $areaApplied^2$ for special-milled) indicates that supply-side considerations influence pricing differently according to rice quality categories.

These results complement earlier studies, especially Qodri et al. (2023), who established an inverse relationship between harvested area and the prices of rice in Indonesia, and Fitrawaty et al. (2023), who showed that rising production volume is linked to low prices of rice in the long term. Like Fahmid et al. (2022), who noted that fertilizer use affects production and hence market prices, our research supports the connection between fertilizer use and rice prices, especially for special-milled rice.

These findings have far-reaching implications for government agencies such as the DA in designing effective rice price stabilization policy and food security programs. The models can forecast future price trends using production drivers to facilitate more responsive interventions in areas with clear pricing trends. Customers can learn more from this study about the factors that influence rice prices by region. Farmers will be able to use these results to better plan fertilizer application and production levels, potentially maximizing their economic return with the added benefit of price stability.

The research was narrowed to major rice-producing areas in Luzon (Regions I to V and CAR) from a four-year timeframe (2020-2023), concentrating only on the connections among harvest area, production volume, fertilizer use, and regional/temporal factors on retail rice prices. Further research would potentially broaden the analysis by adding more variables, including weather conditions, import quantities, and wider economic data to a larger geographical area and a longer period. In spite of these constraints, the results presented here offer useful empirical insights into the determinants of rice prices in the Philippines, providing the basis for more targeted and efficient agricultural and food security policies that take into account the unique needs of various regions and quality segments of rice.

REFERENCES

Antonio, R. J., Valera, H., & Mishra, A. (2024). *Rice price inflation dynamics in the Philippines*. Munich Personal RePEc Archive. https://mpra.ub.uni-muenchen.de/123641/1/MPRA_paper_123641.pdf

Beltran, G. (2024). *Determinants of rice prices in the Philippines*. Determinants of Rice Prices in the Philippines: Recommendations for Policy Reforms to Reduce Supply and Price Volatility. https://doftaxacademy.gov.ph/wp-content/uploads/2024/04/Rice-Research-2-20-2024.pdf

Briones, R., Galang, I. M., & Pastolero, H. L. (2025). *Organic alternatives viable amid rising chemical fertilizer costs in ph, but policy reform is crucial*. PIDS. https://www.pids.gov.ph/details/news/press-releases/or ganic-alternatives-viable-amid-rising-chemical-fertilizer-costs-in-ph-but-policy-reform-is-crucial

Briones, R., Galang, I. M., & Pastolero, H. L. (2024, December). *Prospects for Widespread Adoption of Organic-Based Fertilizers in the Philippines: A Rapid Appraisal*. PIDS. https://pidswebs.pids.gov.ph/CDN/document/pidsdps2430.pdf

Cereals: Retail Prices of Agricultural Commodities by Geolocation, Commodity, Year and Period-PX-Web. (2017). PX-Web. https://openstat.psa.gov.ph/PXWeb/pxweb/en/DB/DB__2M__2018/0042M4ARA01.px/?rxid=c3efa5a4-83bc-4c3c-8c9e-03c6532a5430&fbclid=IwY2xjawI5MFRleHRuA2FlbQIxMAABHepuRhvO4bnuuFCq47zw3dusip9bQMuQwc-UafycznvAj7Np1AGKnX3pFw_aem_zcRNva7uNUzbpjKY3tMOLQ

Fahmid, I. M., Jamil, A., Wahyudi, Agustian, A., Hatta, M., Aldillah, R., Yofa, R. D., Sumedi, Sumaryanto, & Susilowati, S. H. (2022). Study of the impact of increasing the highest retail price of subsidized fertilizer on rice production in Indonesia.

-rice-production-varieties-and-research-what-we-know-so-far/

Open Agriculture, *7*(1), 348–359. https://doi.org/10.1515/opag-2022-0087

Fitrawaty, Hermawan, W., Yusuf, M., & Maipita, I. (2023). A simulation of increasing rice price toward the disparity of income distribution: An evidence from Indonesia. *Heliyon*, *9*(3). https://doi.org/10.1016/j.heliyon.2023.e13785

Gomez, C. J. J. (2024, June 27). *Navigating Challenges and Opportunities: Current Dynamics in the Philippine Rice Industry*. PCAARRD's Industry Strategic Science and Technology Programs. https://ispweb.pcaarrd.dost.gov.ph/navigating-challenges-and-opportunities-current-dynamics-in-the-philippine-rice-industry/#:~:text=According%20to%20ADB%2C%20a%2030,2024%20and%208.45%25%20in%202025.

Office, D. P. (2022, June 3). *From Business Mirror: Philippine rice production, varieties and research: What we know so far*. Official Portal of the Department of Agriculture. https://www.da.gov.ph/from-business-mirror-philippine

Palay and Corn: Area Harvested in Hectares by Ecosystem/Croptype, by Quarter, by Semester, by Region and by Province, 1987-2024-PX-Web. (2017). PX-Web. https://openstat.psa.gov.ph/PXWeb/pxweb/en/DB/DB__2E__CS/0022E4EAHC0.px/?rxid=bdf9d8da-96f1-4100-ae09-18cb3eaeb313%2F&fbclid=IwY2xjawI5L3VleHRuA2FlbQIxMAABHeypq4N0uQTfVdLxd1rLd-vsuaptl9a3BEPucrYBsGnqQNpGqxO_l1OFjw_aem_avKFCzj6yX8dpibI6ro-aA&vcode=27

Palay: Estimated Inorganic Fertilizer Use by Geolocation, Grade and Year by Region, Area Harvested and Year by PHILIPPINES, Area Harvested and Year by PHILIPPINES, FERTILIZER and Year-PX-Web. (2017). PX-Web. https://openstat.psa.gov.ph/PXWeb/pxweb/en/DB/DB__2E__CS/0042E4EPFU0.px/?rxid=bdf9d8da-96f1-4100-ae09-18cb3eaeb313&fbclid=IwY2xjawI5L3hleHRuA2FlbQIxMAABHS6rY4SgU9A4n7NAjX2_fH_GnUwc05pbnNGNOiXGm8I_SRuOXOTyMAeylA_aem_eI61u7FUDjJgdDZF02CkOQ&vcode=Syc4JV

Palay and Corn: Volume of Production in Metric Tons by Ecosystem/Croptype, by Quarter, by Semester, by Region, 1987-2024-PX-Web. (2017). PX-Web. https://openstat.psa.gov.ph/PXWeb/pxweb/en/DB/DB__2E__CS/0112E4EVCP0.px/?rxid=bdf9d8da-96f1-4100-ae09-18cb3eaeb313&fbclid=IwY2xjawI5L2dleHRuA2FlbQIxMAABHTeRNiSYqFCCAMKVyunCj4kU7CP1OXE7xV40Yt2dE0l_SAfJ23NKTwlijw_aem_JIxMy5BFVCU6CY5psrguZw

Qodri, L. A., Ismail, M., Ekawaty, M., & Wahyudi, S. T. (2023). The dynamic impact of rice production and harvested area on agricultural product prices during COVID-19. *Advances in Economics, Business and Management Research*, 275–285. https://doi.org/10.2991/978-2-38476-064-0

Appendix A

Table of Member Contributions

Member	Task
Paul Matthew Gayatiña	Introduction Results: Regular_Milled
Justin Christian Vilbar	Powerpoint Abstract Results: Special_Milled Conclusion
Julia James Berin	Methodology Results: Well_Milled Finalizing

Appendix B

Pearson Correlation Table for Well-milled rice prices

Pearson Correlation Coefficients, N = 24 Prob > |r| under H0: Rho=0				
	wellMilled	**volumePalay**	**areaPalay**	**areaApplied**
wellMilled wellMilled	1.00000	0.30601 0.1459	0.29211 0.1660	0.29528 0.1613
volumePalay volumePalay	0.30601 0.1459	1.00000	0.99268 <.0001	0.99347 <.0001
areaPalay areaPalay	0.29211 0.1660	0.99268 <.0001	1.00000	0.99971 <.0001
areaApplied areaApplied	0.29528 0.1613	0.99347 <.0001	0.99971 <.0001	1.00000

Appendix C

Full-Rank Multiple Linear Regression Model for well-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	42.05102	6.31765	6.66	<.0001
volumePalay	volumePalay	1	-0.00000593	0.00000427	-1.39	0.1899
areaPalay	areaPalay	1	0.00000110	0.00005007	0.02	0.9828
areaApplied	areaApplied	1	0.00002425	0.00004111	0.59	0.5662
region2	region2	1	4.28763	3.23681	1.32	0.2100
region3	region3	1	5.67797	4.33955	1.31	0.2152
region4	region4	1	0.28201	4.60046	0.06	0.9521
region5	region5	1	-0.58488	1.76171	-0.33	0.7456
region6	region6	1	0.93024	4.84208	0.19	0.8509
year2	year2	1	0.20564	0.30830	0.67	0.5174
year3	year3	1	3.78264	0.33978	11.13	<.0001
year4	year4	1	-0.62856	0.37552	-1.67	0.1200

$$\widehat{wellMilled} = 42.05102 - 0.00000593(volumePalay) + 0.00000110(areaPalay) + 0.00002425(areaApplied) + 4.28763(region2) + 5.66797(region3) + 0.28201(\text{region4}) - 0.58488(region5) + 0.93024(region6) + 0.20564(year2) + 3.78264(year3) - 0.62856(year4)$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	105.59367	9.59942	35.04	<.0001
Error	12	3.28712	0.27393		
Corrected Total	23	108.88080			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 35.04
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data

Adjusted R^2

Root MSE	0.52338	R-Square	0.9698
Dependent Mean	43.70792	Adj R-Sq	0.9421
Coeff Var	1.19745		

Adjusted R^2 value: 0.9421

Table C1. Significant Predictors of the Full-Rank MLRM for well-milled rice prices

Variable	p-value	Significant/Not Significant
volumePalay	0.1899	Not Significant
areaPalay	0.9828	Not Significant
areaApplied	0.5662	Not Significant
region2	0.2100	Not Significant
region3	0.2152	Not Significant
region4	0.9521	Not Significant
region5	0.7456	Not Significant
region6	0.8509	Not Significant
year2	0.5174	Not Significant
year3	<0.0001	Significant
year4	0.1200	Not Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.970808	Pr < W	0.6870
Kolmogorov-Smirnov	D	0.14757	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.043454	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.25563	Pr > A-Sq	>0.2500

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.970808
P value: 0.6870
Decision: Fail to reject H_O
Conclusion: Errors are normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : wellresid Residual				
N	Mean	Std Dev	Minimum	Maximum
24	5.62513E-15	0.3780455	-0.6420765	0.9791038

$E[\epsilon_i] = 5.62513E - 15 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of wellresid Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	0.1673	0.0558	1.34	0.2889
Error	20	0.8308	0.0415		

Used year as categorical variable

H_O: Variances of the residuals are constant (Homoscedastic)

H_A: Variances of the residuals are not constant (Heteroscedastic)

Test statistic value: 1.34

P value: 0.2889

Decision: Fail to Reject H_O

Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	2.182
Pr < DW	0.5125
Pr > DW	0.4875
Number of Observations	24
1st Order Autocorrelation	-0.097

H_O: Errors are independent

H_A: Errors are not independent

Test statistic value: 2.182

P value: Pr<Dw 0.5125 Pr>Dw 0.4875

Decision: Fail to reject H_O

Conclusion: Independent

PRESS

Sum of Residuals	0
Sum of Squared Residuals	3.28712
Predicted Residual SS (PRESS)	14.09886

PRESS value: 14.09886

Appendix D

Forward Selection/Backward Elimination/Stepwise Selection Multiple Linear Regression Model for well-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	43.33957	0.26460	163.79	<.0001
volumePalay	volumePalay	1	-0.00000114	1.986824E-7	-5.73	<.0001
region2	region2	1	4.86801	0.47517	10.24	<.0001
region3	region3	1	4.88181	0.60430	8.08	<.0001
year3	year3	1	3.60246	0.26447	13.62	<.0001
year4	year4	1	-0.61373	0.26476	-2.32	0.0324

$$\widehat{wellMilled} = 43.33957 - 0.00000114(volumePalay) + 4.86801(region2) + 4.88181(region3) + 3.60246(year3) - 0.61373(year4)$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	103.84563	20.76913	74.25	<.0001
Error	18	5.03516	0.27973		
Corrected Total	23	108.88080			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 74.25
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data.

Adjusted R^2

Root MSE	0.52890	R-Square	0.9538
Dependent Mean	43.70792	Adj R-Sq	0.9409
Coeff Var	1.21007		

Adjusted R^2 value: 0.9409

Table D1. Significant Predictors of the Automatic Selection MLRM for well-milled rice prices

Variable	p-value	Significant/Not Significant

volumePalay	<0.0001	Significant
region2	<0.0001	Significant
region3	<0.0001	Significant
year3	<0.0001	Significant
year4	0.0324	Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.957348	Pr < W	0.3875
Kolmogorov-Smirnov	D	0.128823	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.0589	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.373025	Pr > A-Sq	>0.2500

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.957348
P value: 0.3875
Decision: Fail to reject H_O
Conclusion: Errors are normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : wellforwardr Residual				
N	Mean	Std Dev	Minimum	Maximum
24	5.62513E-15	0.4678891	-0.7752289	0.8611539

$E[\epsilon_i] = 5.62513E - 15 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of wellforwardr Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	0.2477	0.0826	2.00	0.1458
Error	20	0.8240	0.0412		

Used year as categorical variable
H_O: Variances of the residuals are constant (Homoscedastic)
H_A: Variances of the residuals are not constant (Heteroscedastic)
Test statistic value: 2.00
P value: 0.1458
Decision: Fail to reject H_O
Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	2.356
Pr < DW	0.7240
Pr > DW	0.2760
Number of Observations	24
1st Order Autocorrelation	-0.233

H_O: Errors are independent
H_A: Errors are not independent
Test statistic value: 2.356
P value: Pr<Dw 0.7240 Pr>Dw 0.2760
Decision: Fail to reject H_O
Conclusion: Errors are independent

PRESS

Sum of Residuals	0
Sum of Squared Residuals	5.03516
Predicted Residual SS (PRESS)	9.34055

PRESS value: 9.34055

Appendix E

Comparison of different MLRM models for well-milled rice prices

Table E1. Comparison of different multiple linear regression models for well-milled rice prices

Characteristics	Full Rank	Forward/Backward/Stepwise
Variables	volumePalay areaPalay areaApplied region2 region3 region4 region5 region6 year2 year3 year4	volumePalay Region2 Region3 Year3 year4
Significance	Good Fit	Good Fit
AdjRsq	0.9421	0.9409
Significant Predictors	year3	volumePalay Region2 Region3 Year3 year4
Not Significant Predictors	volumePalay areaPalay areaApplied Region2 Region3 Region4 region5 Region6 Year2 year4	
Normality	Normal	Normal
Proc Means	0	0
Homoscedasticity	Homo	Homo
Independence	Independent	Independent
PRESS	14.09886	9.34055
MSE	0.27393	0.27973
C(p)	p=11 12.0000	p=5 6.3814

AIC	-23.7130	-25.4786
SBC	-9.57634	-18.41026
Multicollinearity	Multicollinear	Not Multicollinear

Appendix F

Pearson Correlation Table for Regular-milled rice prices

Pearson Correlation Coefficients, N = 24 Prob > |r| under H0: Rho=0				
	regularMilled	volumePalay	areaPalay	areaApplied
regularMilled regularMilled	1.00000	-0.30323 0.1498	-0.32651 0.1194	-0.32743 0.1183
volumePalay volumePalay	-0.30323 0.1498	1.00000	0.99268 <.0001	0.99347 <.0001
areaPalay areaPalay	-0.32651 0.1194	0.99268 <.0001	1.00000	0.99971 <.0001
areaApplied areaApplied	-0.32743 0.1183	0.99347 <.0001	0.99971 <.0001	1.00000

Appendix G

Full-Rank Multiple Linear Regression Model for Regular-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	83093	17873	4.65	0.0006
volumePalay	volumePalay	1	0.01831	0.01207	1.52	0.1551
areaPalay	areaPalay	1	-0.36429	0.14165	-2.57	0.0245
areaApplied	areaApplied	1	0.19051	0.11630	1.64	0.1273
region2	region2	1	14485	9157.23735	1.58	0.1397
region3	region3	1	27151	12277	2.21	0.0472
region4	region4	1	-14765	13015	-1.13	0.2788
region5	region5	1	12272	4984.04005	2.46	0.0299
region6	region6	1	-12961	13699	-0.95	0.3627
year2	year2	1	2414.14235	872.21666	2.77	0.0170
year3	year3	1	18135	961.26711	18.87	<.0001
year4	year4	1	268.91986	1062.38872	0.25	0.8045

$$\widehat{regularMilled}^{3} = 83093 + 0.01831(volumePalay) - 0.36429(areaPalay) + 0.19051(areaApplied) + 14485(region2) + 27151(region3) - 14765(region4) + 12272(region5) - 12961(region6) + 2414.14235(year2) + 18135(year3) + 268.91986(year4)$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	2178878368	198079852	90.35	<.0001
Error	12	26309388	2192449		
Corrected Total	23	2205187756			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 90.35
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data

Adjusted R^2

Root MSE	1480.69206	R-Square	0.9881
Dependent Mean	57268	Adj R-Sq	0.9771
Coeff Var	2.58556		

Adjusted R^2 value: 0.9771

Table G1. Significant Predictors of the Full-Rank MLRM for regular-milled rice prices

Variable	p-value	Significant/Not Significant
volumePalay	0.1551	Not Significant
areaPalay	0.0245	Significant
areaApplied	0.1273	Not Significant
region2	0.1397	Not Significant
region3	0.0472	Significant
region4	0.2788	Not Significant
region5	0.0299	Significant
region6	0.3627	Not Significant
year2	0.0170	Significant
year3	<0.0001	Significant
year4	0.8045	Not Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.955465	Pr < W	0.3542
Kolmogorov-Smirnov	D	0.117123	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.0486	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.344601	Pr > A-Sq	>0.2500

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.955465
P value: 0.3542
Decision: Fail to reject H_O
Conclusion: Errors are normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : regresid Residual				
N	Mean	Std Dev	Minimum	Maximum
24	-2.24342E-10	1069.53	-2087.31	1621.11

$E[\epsilon_i] =- 2.24342E - 10 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of regresid Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	3.664E12	1.221E12	0.96	0.4290
Error	20	2.534E13	1.267E12		

Used year as categorical variable

H_O: Variances of the residuals are constant (Homoscedastic)

H_A: Variances of the residuals are not constant (Heteroscedastic)

Test statistic value: 0.96

P value: 0.4290

Decision: Fail to reject H_O

Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	2.158
Pr < DW	0.4907
Pr > DW	0.5093
Number of Observations	24
1st Order Autocorrelation	-0.146

H_O: Errors are independent

H_A: Errors are not independent

Test statistic value: 2.158

P value: Pr<Dw 0.4907 Pr>Dw 0.5093

Decision: Fail to reject H_O

Conclusion: Errors are independent

PRESS

Sum of Residuals	-4.66504E-9
Sum of Squared Residuals	26309388
Predicted Residual SS (PRESS)	128139066

Press value: 128,139,066

Appendix H

Forward Selection/Stepwise Selection Multiple Linear Regression Model for Regular-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	54395	964.33604	56.41	<.0001
region2	region2	1	-10465	2089.66350	-5.01	<.0001
year3	year3	1	18466	1798.49599	10.27	<.0001

$$\widehat{regularMilled}^{3} = 54395 - 10465(region2) + 18466(year3)$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1899519207	949759603	65.25	<.0001
Error	21	305668549	14555645		
Corrected Total	23	2205187756			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 65.25
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data

Adjusted R^2

Root MSE	3815.18613	R-Square	0.8614
Dependent Mean	57268	Adj R-Sq	0.8482
Coeff Var	6.66200		

Adjusted R^2 value: 0.8482

Table H1. Significant Predictors of the Forward Selection MLRM for regular-milled rice prices

Variable	p-value	Significant/Not Significant

region2	<0.0001	Significant
year3	<0.0001	Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.936867	Pr < W	0.1389
Kolmogorov-Smirnov	D	0.161687	Pr > D	0.1005
Cramer-von Mises	W-Sq	0.120731	Pr > W-Sq	0.0563
Anderson-Darling	A-Sq	0.652886	Pr > A-Sq	0.0815

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.936867
P value: 0.1389
Decision: Fail to reject H_O
Conclusion: Errors are normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : regforwardr Residual				
N	Mean	Std Dev	Minimum	Maximum
24	6.972793E-12	3645.54	-8187.17	5840.85

$E[\epsilon_i] = 6.972793E - 12 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of regforwardr Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	2.775E14	9.249E13	0.31	0.8203
Error	20	6.032E15	3.016E14		

Used year as categorical variable
H_O: Variances of the residuals are constant (Homoscedastic)
H_A: Variances of the residuals are not constant (Heteroscedastic)
Test statistic value: 0.31
P value: 0.8203
Decision: Fail to reject H_O
Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	2.133
Pr < DW	0.5636
Pr > DW	0.4364
Number of Observations	24
1st Order Autocorrelation	-0.107

H_O: Errors are independent
H_A: Errors are not independent
Test statistic value: 2.133
P value: Pr<Dw 0.5636 Pr>Dw 0.4364
Decision: Fail to reject H_O
Conclusion: Errors are independent

PRESS

Sum of Residuals	0
Sum of Squared Residuals	305668549
Predicted Residual SS (PRESS)	383380168

Press value: 383,380,168

Appendix I

Backward Elimination Multiple Linear Regression Model for Regular-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	58113	761.86857	76.28	<.0001
areaPalay	areaPalay	1	-0.02533	0.00183	-13.83	<.0001
region3	region3	1	12976	1153.38626	11.25	<.0001
region5	region5	1	6667.79849	890.98859	7.48	<.0001
year2	year2	1	2744.50486	794.38642	3.45	0.0028
year3	year3	1	19376	794.38649	24.39	<.0001

$$\widehat{regularMilled}^{3} = 58113 - 0.02533(areaPalay) + 12976(region3) + 6667.79849(region5) + 2744.50486(year2) + 19376(year3)$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2159752172	431950434	171.12	<.0001
Error	18	45435584	2524199		
Corrected Total	23	2205187756			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 171.12
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data

Adjusted R^2

Root MSE	1588.77283	R-Square	0.9794
Dependent Mean	57268	Adj R-Sq	0.9737
Coeff Var	2.77428		

Adjusted R^2 value: 0.9737

Table I1. Significant Predictors of the Forward Selection MLRM for regular-milled rice prices

Variable	p-value	Significant/Not Significant
areaPalay	<0.0001	Significant
region3	<0.0001	Significant

region5	<0.0001	Significant
year2	0.0028	Significant
year3	<0.0001	Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.965128	Pr < W	0.5497
Kolmogorov-Smirnov	D	0.130448	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.073033	Pr > W-Sq	0.2467
Anderson-Darling	A-Sq	0.389779	Pr > A-Sq	>0.2500

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.965128
P value: 0.5497
Decision: Fail to reject H_O
Conclusion: Errors are normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : regbackwardr Residual				
N	Mean	Std Dev	Minimum	Maximum
24	2.425319E-12	1405.51	-2465.34	3199.98

$E[\epsilon_i] = 2.425319E - 12 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of regbackwardr Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	4.458E13	1.486E13	3.02	0.0540
Error	20	9.853E13	4.926E12		

Used year as categorical variable
H_O: Variances of the residuals are constant (Homoscedastic)
H_A: Variances of the residuals are not constant (Heteroscedastic)
Test statistic value: 3.02
P value: 0.0540
Decision: Fail to reject H_O
Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	2.558
Pr < DW	0.8320
Pr > DW	0.1680
Number of Observations	24
1st Order Autocorrelation	-0.331

H_O: Errors are independent
H_A: Errors are not independent
Test statistic value: 2.558
P value: Pr<Dw 0.8320 Pr>Dw 0.1680
Decision: Fail to reject H_O
Conclusion: Errors are independent

PRESS

Sum of Residuals	0
Sum of Squared Residuals	45435584
Predicted Residual SS (PRESS)	90785900

Press value: 90,785,900

Appendix J

Table Comparison of different MLRM models for Regular-milled rice prices

	Full Rank	Forward/Stepwise	Backward
Variables	volumePalay areaPalay areaApplied region2 region3 region4 region5 region6 year2 year3 year4	Region2 year3	areaPalay Region3 Region5 Year2 year3
Significance	Good Fit	Good Fit	Good Fit
AdjRsq	0.9771	0.8482	0.9737
Significant Predictors	areaPalay Region5 Year2 year3	Region2 year3	areaPalay Region3 Region5 Year2 year3
Not Significant Predictors	volumePalay		

	areaApplied Region2 Region3 Region4 Region6 year4		
Normality	Normal	Normal	Normal
Proc Means	0	0	0
Homoscedasticity	Homo	Homo	Homo
Independence	Independent	Independent	Independent
PRESS	128,139,066	383,380,168	90,785,900
MSE	2,192,449	14,555,645	2,524,199
C(p)	p=11 12.0000	p=2 121.4188	p=5 8.7237
AIC	357.7772	398.6390	358.8901
SBC	371.91383	402.17316	365.95838
Multicollinearity	Multicollinear	Not Multicollinear	Not Multicollinear

Appendix K

Pearson Correlation Table for Special-milled rice prices

Pearson Correlation Coefficients, N = 24 Prob > |r| under H0: Rho=0				
	specialMilled	volumePalay	areaPalay	areaApplied
specialMilled specialMilled	1.00000	0.60678 0.0017	0.56821 0.0038	0.57978 0.0030
volumePalay volumePalay	0.60678 0.0017	1.00000	0.99268 <.0001	0.99347 <.0001
areaPalay areaPalay	0.56821 0.0038	0.99268 <.0001	1.00000	0.99971 <.0001
areaApplied areaApplied	0.57978 0.0030	0.99347 <.0001	0.99971 <.0001	1.00000

Appendix L

Full-Rank Multiple Linear Regression Model for Special-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	5.12141	0.56568	9.05	<.0001
volumePalay	volumePalay	1	4.842029E-8	1.509858E-7	0.32	0.7540
areaPalay	areaPalay	1	-0.00000456	0.00000239	-1.91	0.0808
areaApplied2		1	3.48215E-12	1.62982E-12	2.14	0.0539
region2	region2	1	0.19656	0.12246	1.61	0.1345
region3	region3	1	0.17949	0.14270	1.26	0.2324
region4	region4	1	-0.75701	0.35895	-2.11	0.0566
region5	region5	1	-0.09977	0.04782	-2.09	0.0590
region6	region6	1	-0.82738	0.37015	-2.24	0.0452
year2	year2	1	0.00945	0.01052	0.90	0.3867
year3	year3	1	0.06478	0.01182	5.48	0.0001
year4	year4	1	-0.01415	0.01246	-1.14	0.2781

$$\begin{aligned} ln(\widehat{specialMilled}) &= 5.12141 + (4.842029E - 8)(volumePalay) - 0.00000456(areaPalay) + 3.48215E - 12(areaApplied)^2 \\ &+ 0.19656(region2) + 0.17949(region3) - 0.75701(region4) - 0.09977(region5) - 0.82738(region6) + 0.00945(year2) \\ &+ 0.06478(year3) - 0.01415(year4) \end{aligned}$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	0.06167	0.00561	17.37	<.0001
Error	12	0.00387	0.00032281		
Corrected Total	23	0.06555			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 17.37
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data

Adjusted R^2

Root MSE	0.01797	R-Square	0.9409
Dependent Mean	3.94319	Adj R-Sq	0.8867
Coeff Var	0.45564		

Adjusted R^2 value: 0.8867

Table G1. Significant Predictors of the Full-Rank MLRM for regular-milled rice prices

Variable	p-value	Significant/Not Significant
volumePalay	0.7540	Not Significant
areaPalay	0.0808	Not Significant
areaApplied2	0.0539	Not Significant
region2	0.1345	Not Significant
region3	0.2324	Not Significant
region4	0.0566	Not Significant
region5	0.0590	Not Significant
region6	0.0452	Significant
year2	0.3867	Not Significant
year3	0.0001	Significant
year4	0.2781	Not Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.941549	Pr < W	0.1766
Kolmogorov-Smirnov	D	0.128136	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.068813	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.455439	Pr > A-Sq	0.2484

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.941549
P value: 0.1766
Decision: Fail to reject H_O
Conclusion: Errors are normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : specialresid Residual				
N	Mean	Std Dev	Minimum	Maximum
24	1.276756E-15	0.0129778	-0.0376662	0.0227042

$E[\epsilon_i] = 1.276756E - 15 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of specialresid Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	4.576E-7	1.525E-7	1.86	0.1691
Error	20	1.641E-6	8.205E-8		

Used year as categorical variable

H_O: Variances of the residuals are constant (Homoscedastic)

H_A: Variances of the residuals are not constant (Heteroscedastic)

Test statistic value: 1.86

P value: 0.1691

Decision: Fail to reject H_O

Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	1.718
Pr < DW	0.1600
Pr > DW	0.8400
Number of Observations	24
1st Order Autocorrelation	0.043

H_O: Errors are independent

H_A: Errors are not independent

Test statistic value: 1.718

P value: Pr<Dw 0.1600 Pr>Dw 0.8400

Decision: Fail to reject H_O

Conclusion: Errors are independent

PRESS

Sum of Residuals	2.83031E-14
Sum of Squared Residuals	0.00387
Predicted Residual SS (PRESS)	0.01618

Press value: 0.01618

Appendix M

Forward Selection//Stepwise Selection Multiple Linear Regression Model for Special-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	3.88312	0.00866	448.34	<.0001
areaApplied2		1	2.13893E-13	2.55885E-14	8.36	<.0001
region4	region4	1	0.04634	0.01266	3.66	0.0017
year3	year3	1	0.06637	0.01033	6.42	<.0001
year4	year4	1	-0.02362	0.01033	-2.29	0.0339

$$ln(\widehat{specialMilled}) = 3.88312 + (2.13893E - 13)(areaApplied)^2 + 0.04634(region4) + 0.06637(year3) - 0.02362(year4)$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	0.05744	0.01436	33.64	<.0001
Error	19	0.00811	0.00042687		
Corrected Total	23	0.06555			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 33.64
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data

Adjusted R^2

Root MSE	0.02066	R-Square	0.8763
Dependent Mean	3.94319	Adj R-Sq	0.8502
Coeff Var	0.52396		

Adjusted R^2 value: 0.8502

Table G1. Significant Predictors of the Full-Rank MLRM for regular-milled rice prices

Variable	p-value	Significant/Not Significant
areaApplied2	<.0001	Significant
region4	0.0017	Significant
year3	<.0001	Significant
year4	0.0339	Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.813482	Pr < W	0.0005
Kolmogorov-Smirnov	D	0.173949	Pr > D	0.0595
Cramer-von Mises	W-Sq	0.197519	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.243964	Pr > A-Sq	<0.0050

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.813482
P value: 0.0005
Decision: Reject H_O
Conclusion: Errors are not normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : specialforwardr Residual				
N	Mean	Std Dev	Minimum	Maximum
24	1.184238E-15	0.0187785	-0.0680497	0.0281903

$E[\epsilon_i] = 1.184238EE - 15 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of specialforwardr Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	3.55E-6	1.183E-6	1.42	0.2652
Error	20	0.000017	8.309E-7		

Used year as categorical variable
H_O: Variances of the residuals are constant (Homoscedastic)
H_A: Variances of the residuals are not constant (Heteroscedastic)
Test statistic value: 1.42
P value: 0.2652

Decision: Fail to reject H_O
Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	1.578
Pr < DW	0.0593
Pr > DW	0.9407
Number of Observations	24
1st Order Autocorrelation	0.185

H_O: Errors are independent
H_A: Errors are not independent
Test statistic value: 1.578
P value: Pr<Dw 0.0593 Pr>Dw 0.9407
Decision: Fail to reject H_O
Conclusion: Errors are independent

PRESS

Sum of Residuals	2.04769E-14
Sum of Squared Residuals	0.00811
Predicted Residual SS (PRESS)	0.01269

Press value: 0.01269

Appendix N

Backward Elimination Multiple Linear Regression Model for Special-milled rice prices

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	3.92427	0.00903	434.59	<.0001
areaApplied2		1	1.16602E-13	2.50498E-14	4.65	0.0002
region5	region5	1	-0.04667	0.01126	-4.14	0.0006
region6	region6	1	-0.03885	0.01237	-3.14	0.0057
year3	year3	1	0.06625	0.00947	7.00	<.0001
year4	year4	1	-0.02442	0.00947	-2.58	0.0189

$$\widehat{ln(specialMilled)} = 3.92427 + (1.16602E - 13)(areaApplied)^2 - 0.04667(region5) - 0.03885(region6) + 0.06625(year3) - 0.02442(year4)$$

Test for Significance of the Model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	0.05910	0.01182	32.98	<.0001
Error	18	0.00645	0.00035839		
Corrected Total	23	0.06555			

H_O: The model is not a good fit for the data
H_A: The model is a good fit for the data
Test statistic value: 32.98
P-value: <0.0001
Decision: Reject H_O
Conclusion: The model is a good fit for the data

Adjusted R^2

Root MSE	0.01893	R-Square	0.9016
Dependent Mean	3.94319	Adj R-Sq	0.8742
Coeff Var	0.48010		

Adjusted R^2 value: 0.8742

Table G1. Significant Predictors of the Full-Rank MLRM for regular-milled rice prices

Variable	p-value	Significant/Not Significant
areaApplied2	0.0002	Significant
region5	0.0006	Significant
region6	0.0057	Significant
year3	<.0001	Significant
year4	0.0189	Significant

Test for Normality

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.920937	Pr < W	0.0613
Kolmogorov-Smirnov	D	0.165838	Pr > D	0.0865
Cramer-von Mises	W-Sq	0.103492	Pr > W-Sq	0.0963
Anderson-Darling	A-Sq	0.623103	Pr > A-Sq	0.0942

H_O: Errors are normally distributed
H_A: Errors are not normally distributed
Test statistic value: 0.920937
P value: 0.0613
Decision: Fail to reject H_O
Conclusion: Errors are normally distributed

Mean of the Probability Distribution of the Residuals

Analysis Variable : specialbackwardr Residual				
N	Mean	Std Dev	Minimum	Maximum
24	4.625929E-16	0.0167474	-0.0515905	0.0314122

$E[\epsilon_i] = 4.625929E - 16 = 0$

Test for Homoscedasticity

Levene's Test for Homogeneity of specialbackwardr Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
year	3	8.058E-7	2.686E-7	0.85	0.4823
Error	20	6.311E-6	3.156E-7		

Used year as categorical variable
H_O: Variances of the residuals are constant (Homoscedastic)
H_A: Variances of the residuals are not constant (Heteroscedastic)
Test statistic value: 0.85
P value: 0.4823

Decision: Fail to reject H_O
Conclusion: Homoscedastic

Test for Independence

Durbin-Watson D	1.490
Pr < DW	0.0521
Pr > DW	0.9479
Number of Observations	24
1st Order Autocorrelation	0.165

H_O: Errors are independent
H_A: Errors are not independent
Test statistic value: 1.490
P value: Pr<Dw 0.0521 Pr>Dw 0.9479
Decision: Fail to reject H_O
Conclusion: Errors are independent

PRESS

Sum of Residuals	0
Sum of Squared Residuals	0.00645
Predicted Residual SS (PRESS)	0.01436

Press value: 0.01436

Appendix O

Table Comparison of different models for Special-milled rice prices

	Full Rank	Forward/Stepwise	Backward
Variables	volumePalay areaPalay areaApplied2 region2 region3 region4 region5 region6 year2 year3 year4	areaApplied2 region4 year3 year4	areaApplied2 region5 region6 year3 year4
Significance	Good Fit	Good Fit	Good Fit
AdjRsq	0.8867	0.8502	0.8742
Significant Predictors	region6 year3	areaApplied2 region4 year3	areaApplied2 region5 region6

		year4	year3 year4
Not Significant Predictors	volumePalay areaPalay areaApplied region2 region3 region4 region5 year2 year4		
Normality	Normal	Not Normal	Normal
Proc Means	0	0	0
Homoscedasticity	Homo	Homo	Homo
Independence	Independent	Independent	Independent
PRESS	0.01618	0.01269	0.01436
MSE	0.00032281	0.00042687	0.00035839
C(p)	p=11 12.0000	p=4 11.1248	p=5 7.9838
AIC	-185.5583	-181.8235	-185.3179
SBC	-171.42164	-175.93327	-178.24959
Multicollinearity	Multicollinear	Not Multicollinear	Not Multicollinear

Appendix P

SAS Codes for the Research Proper

```
proc print data = rice;
run;

proc corr data=rice;
   var wellMilled volumePalay areaPalay areaApplied;
run;

proc reg data = rice;
model wellMilled = volumePalay areaPalay areaApplied region2 region3 region4 region5  region6
year2 year3 year4
            / selection = rsquare adjrsq mse cp AIC SBC press alpha = 0.05;
run;

proc reg data = rice;
model wellMilled = volumePalay areaPalay areaApplied region2 region3 region4 region5  region6
year2 year3 year4/
clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = wellresult p = wellyhat r = wellresid;
run;

proc univariate data = wellresult freq plot normal;
var wellresid;
run;

proc means data = wellresult ;
var wellresid;
run;

proc anova data = wellresult;
class year;
model wellresid = year;
means year / hovtest=Levene;
run;
```

```
proc reg data = rice;
model wellMilled = volumePalay areaPalay areaApplied region2 region3 region4 region5  region6
year2 year3 year4
            / selection = forward slentry = 0.05
            clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = wellforward  p = wellforwardyhat r = wellforwardr ;
run;

proc univariate data = wellforward freq plot normal;
var wellforwardr;
run;

proc means data = wellforward ;
var wellforwardr;
run;

proc anova data = wellforward;
class year;
model wellforwardr = year;
means year / hovtest=Levene;
run;

proc reg data = rice;
model wellMilled = volumePalay areaPalay areaApplied region2 region3 region4 region5  region6
year2 year3 year4
            / selection = backward slstay = 0.05
            clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = wellbackward  p = wellbackwardyhat r = wellbackwardr ;
run;

proc reg data = rice;
model wellMilled = volumePalay areaPalay areaApplied region2 region3 region4 region5  region6
year2 year3 year4
            / selection = stepwise slentry = 0.05 slstay = 0.05
            clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = wellstep  p = wellstepyhat r = wellstepr ;
run;

proc corr data=rice;
   var regularMilled volumePalay areaPalay areaApplied;
run;
```

```
data rice2;
set rice;
regularMilled3 = regularMilled*regularMilled*regularMilled;
proc print data = rice2;
run;

proc reg data = rice2;
model regularMilled3 = volumePalay areaPalay areaApplied region2 region3 region4 region5
region6 year2 year3 year4
        / selection = rsquare adjrsq mse cp AIC SBC press alpha = 0.05;
run;

proc reg data = rice2;
model regularMilled3 = volumePalay areaPalay areaApplied region2 region3 region4 region5
region6 year2 year3 year4/
clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = regresult p = regyhat r = regresid;
run;

proc univariate data = regresult freq plot normal;
var regresid;
run;

proc means data = regresult ;
var regresid;
run;

proc anova data = regresult;
class year;
model regresid = year;
means year / hovtest=Levene;
run;

proc reg data = rice2;
model regularMilled3 = volumePalay areaPalay areaApplied region2 region3 region4 region5
region6 year2 year3 year4
        / selection = forward slentry = 0.05
        clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = regforward  p = regforwardyhat r = regforwardr ;
run;
```

```
proc univariate data = regforward freq plot normal;
var regforwardr;
run;

proc means data = regforward;
var regforwardr;
run;

proc anova data = regforward;
class year;
model regforwardr = year;
means year / hovtest=Levene;
run;

proc reg data = rice2;
model regularMilled3 = volumePalay areaPalay areaApplied region2 region3 region4 region5
region6 year2 year3 year4
        / selection = backward slstay = 0.05
        clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = regbackward  p = regbackwardyhat r = regbackwardr ;
run;

proc univariate data = regbackward freq plot normal;
var regbackwardr;
run;

proc means data = regbackward;
var regbackwardr;
run;

proc anova data = regbackward;
class year;
model regbackwardr = year;
means year / hovtest=Levene;
run;
```

```
proc reg data = rice2;
model regularMilled3 = volumePalay areaPalay areaApplied region2 region3 region4 region5
region6 year2 year3 year4
        / selection = stepwise slentry = 0.05 slstay = 0.05
        clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = regstep  p = regstepyhat r = regstepr ;
run;

proc corr data=rice;
   var specialMilled volumePalay areaPalay areaApplied;
run;

data rice3;
set rice;
areaApplied2=areaApplied*areaApplied;
lnspecialMilled=log(specialMilled);
proc print data = rice3;
run;

proc reg data = rice3;
model lnspecialMilled = volumePalay areaPalay areaApplied2 region2 region3 region4 region5
region6 year2 year3 year4
        / selection = rsquare adjrsq mse cp AIC SBC press alpha = 0.05;
run;

proc reg data = rice3;
model lnspecialMilled = volumePalay areaPalay areaApplied2 region2 region3 region4 region5
region6 year2 year3 year4/
clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = specialresult p = specialyhat r = specialresid;
run;

proc univariate data = specialresult freq plot normal;
var specialresid;
run;

proc means data = specialresult;
var specialresid;
run;
```

```
proc anova data = specialresult;
class year;
model specialresid = year;
means year / hovtest=Levene;
run;

proc reg data = rice3;
model lnspecialMilled = volumePalay areaPalay areaApplied2 region2 region3 region4 region5
region6 year2 year3 year4
        / selection = forward slentry = 0.05
        clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = specialforward  p = specialforwardyhat r = specialforwardr ;
run;

proc univariate data = specialforward freq plot normal;
var specialforwardr;
run;

proc means data = specialforward;
var specialforwardr;
run;

proc anova data = specialforward;
class year;
model specialforwardr = year;
means year / hovtest=Levene;
run;

proc reg data = rice3;
model lnspecialMilled = volumePalay areaPalay areaApplied2 region2 region3 region4 region5
region6 year2 year3 year4
        / selection = backward slstay = 0.05
        clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = specialbackward  p = specialbackwardyhat r = specialbackwardr ;
run;

proc univariate data = specialbackward freq plot normal;
var specialbackwardr;
run;
```

```
            proc means data = specialbackward;
var specialbackwardr;
run;

proc anova data = specialbackward;
class year;
model specialbackwardr = year;
means year / hovtest=Levene;
run;

proc reg data = rice3;
model lnspecialMilled = volumePalay areaPalay areaApplied2 region2 region3 region4 region5
region6 year2 year3 year4
             / selection = stepwise slentry = 0.05 slstay = 0.05
            clm cli clb covb corrb all r influence dwprob vif tol collin alpha = .05;
output out = specialstep  p = specialstepyhat r = specialstepr ;
run;
```