ABSTRACT

SHEN, ZHANGXIN. Modeling Farm-Retail Price Spread in the U.S. Pork Industry. (Under the direction of Michael K. Wohlgenant.)

The farm-retail price spread is the difference between the retail price of a product and its farm value. It changes with changes in factor prices, the efficiency of providing services, and the quantity and quality of services embodied in the final product. A model was derived by Box-Cox transform base on the relative price spread mode for the U.S. pork industry. The new model analyses the determinant of margins more accurately. The results indicate the log of farm-retail price spread is significantly and positively related to increases in log of retail price and log of quantity of farm input. And the relationship between the price spread and industry costs is indeterminate. The results point to the strong possibility of spurious correlation between the price spread and concentration variables. It suggests other possibly unobserved variables correlated with trend are spuriously indicating concentration ratio, has a significant effect on the price spread. Another major implication of this study is that variables used on the regression need to be detrended in estimation.

Modeling Farm-Retail Price Spread in the U.S. Pork Industry

by
Zhangxin Shen

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Arts

Economics Co-major in Statistics

Raleigh, North Carolina

2010

APPROVED BY:

Michael K. Wohlgenant Committee Chair	Howard D. Bondell Committee Co-Chair

Kelly D.Zering

DEDICATION

To my family

Without their constant love and encouragement, I would not complete my education successfully.

BIOGRAPHY

Zhangxin Shen was born in Huzhou, Zhejiang, China in 1986.

He received his Bachelor degree in Applied Biology from Zhejiang University.

Upon graduation in 2008, he enrolled in the Master program in the Department of Economics and Department of Statistics at the North Carolina State University.

ACKNOWLEDGMENTS

I would like to thank Pro. Michael K. Wohlgenant for providing excellent guidance and professional suggestions and helping me get on track within a very tight schedule.

I am very grateful to Dr. Howard D.Bondell for assisting me with the statistical questions.

Finally, I want to thank you Pro.Kelly D. Zering for his great comments on my thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

In the simplest form of price theory, we assume that the consumers and producers meet directly. In real life, however, the price is different between the producers and final consumers. The marketing margin is characterized as some function of the difference between retail and farm price of a given farm product. A farm-retail price spread (marketing margin or marketing charge) "is the difference between the retail price of a product and its farm value ----the payment (adjusted for by-product values) to farmers for an equivalent quantity of farm products" (U.S. Department of Agriculture, 1957, p.1). The marketing margin represents payments for all assembling, processing, transporting, and retailing charges added to farm products (Elitzak, 1996), which can be separated into the costs incurred and profits earned by all agencies. Margins vary greatly among commodities (George and King). It changes with changes in factor prices, the efficiency of providing services, and the quantity and quality of services embodied in the final product (Tomek and Robinson).

We can measure the cost of providing marketing services by the marketing margin. Producers can meet the consumer expectations and market their products through the price spread. Also they can use the price spread to measure the efficiency and equity of the food marketing system.

Another interesting thing about marketing margins is that "high and increasing price spreads often lead to controversy. Livestock producers often blame low livestock prices on high price spreads. And consumers blame high retail prices on high price spreads. Increasing price spreads can both inflate retail prices and deflate farm prices" (Hahn).

A lot of questions have been asked about the margins, such as what will cause different marketing margins? How does this difference happen? How have margins changed over time? Are margins determined via markup pricing? (Wohlgenant, 2001)

So far, substantial research has been done on the questions related to price differences. This thesis will focus on the farm-retail price spread for pork, which is one of the most important items in the consumer's food budget in the US. The main objective is to find a model which can analyze the determinants of margins accurately. In the study, we extend the relative price spread model (Wohlgenant and Mullen) and the markup pricing model (George and King) by including two new variables, four-firm concentration and time trend. These models for the US pork industry are estimated by the Box-Cox Method, and the method of dummy variables is used to account for the structural change events.

2 Literature Review

In many studies on price spreads (e.g.Dalrymple,1961), it is assumed that price spreads are determined in one of the three ways: Constant percentage spread, Absolute spreads, Price spread and quantity handled may have certain relationship. George and King pointed out that it seems appropriate to assume that the price spreads are determined as a combination of percentage and absolute margins. And Waugh also mentioned that many studies suggested that the price spreads are neither constant percentage nor constant absolute amounts, but somewhere in between the two. Under this assumption, the marketing margin is specified as a linear function of retail price and marketing input prices.

As discussed by Gardner (1975), the marketing margin can be measured by the difference between the retail and farm price, by the ratio of the prices, by the farmer's share of the food dollar, or by the percentage marketing margin. Gardner (1975) focuses on the retail-farm price ratio, the closely related percentage margin, and the farmer's share of retail food expenditures. Gardner found out that one implication of the results is that no simple markup pricing rule----a fixed percentage margin, a fixed absolute margin, or a combination of the two----can in general accurately depict the relationship between the farm and retail price. So he developed implications for retail-farm price ratio and farmer's share of the retail dollar using a two-factor, single product long-run competitive equilibrium model.

Later, in 1987, Wohlgenant and Mullen derived a new model -----the relative price spread model. In the model, relating the price spread to industry output and marketing input prices where the price spread and marketing inputs are deflated by retail price allows simultaneous changes in demand and supply conditions.

In contrast to the markup pricing model, the relative price spread model indicates that there is no fixed relationship between the price spreads and retail price (Wohlgenant and Mullen). Wohlgenant and Mullen used nonnested econometric testing procedures to test the relative price spread model and the markup pricing model by George and King. The test results indicate rejection of the markup pricing model compared to the relative price spread model.

Capps, Byrnes, and Williams further developed the relative price model in an application to Lamb industry. The relative price model was first augmented to account for three unique factors indigenous to lamb industry, the four-firm concentration ratio for lamb packing industry, a slope shifter for four-firm concentration ratio for 1986 through 1990, and bimonthly consumption of lamb per capita.

3 Data

Table A in the Appendix contains the definitions of the main variables and the symbol used to represent the variables in the model. Q is per capita quantity of pork disappearance (million pounds of pork, carcass weight, divided by civilian population in millions). IC is an index of marketing costs for pork, 1982 is taken as 100. In this paper, data for IC is from producer price index of fuels and related products and power (PPI) and index of earnings of employees in animal slaughtering and processing[1]. The weights that we use are 0.6 for index of earnings and 0.4 for PPI (Wohlgenant, 2010). Four-Firm Concentration Ratio measures the total market share of the 4 largest firms in an industry. Data for the four firm concentration ratios (CR_4) for pork were obtained from the USDA, Packers and Stockyards Statistical Reports[2]. As shown in Figure 1, the four-firm concentration ratio for pork slaughter indicates that, over time, the top four firms are accounting for a growing share of the overall market. The ratio from 1970 to 2008 rose from 31.6% to 64.5%, but it remained relatively stable from 1970 to 1980.

All price-related variables (M, P_f, P_r and IC) in the table are real, deflated by the US consumer price index. Data for the margin M, farm price P_f (which is equal to gross farm value subtracting the by-product value), retail price P_r and Quantity Q were all obtained from the Economics Research Service (ERS) of the USDA. All of the collected data are from 1970 to 2008, a total of 39 observations.

Figure 2 is the Farm-Retail price spread VS year, 1970-2008, for pork. From the figure, there exists rapid growth in 1998 and 1999, and a rapid decrease in 1973.

[1] The data of index and CPI from USDL, http://www.bls.gov/data/. And the index of earnings of employees in animal slaughtering and processing was only published from 1976 to 2010, a short program was written in SAS to estimate the data in 1970-1975.

[2] The CR_4 are from USDA, *Packers and Stockyards Statistical Reports*. Since the GIPSA program started in 1980, the CR_4 before 1970 to 1979 were estimated by the trend from the figure: Four-Firm Concentration Rations: Cattle and Hog Sectors, 1963-2006 in livestock Marketing and Competition Issues.

[3] The quantity is from USDA, http://www.ers.usda.gov/Search/?qt=mtredsu and http://www.ers.usda.gov/Search/?qt=%27Table+10%E2%80%94U.S.+Meat+Supply+%26+Use .

The price spread and retail price, farm price are from USDA, http://www.ers.usda.gov/Data/MeatPriceSpreads/ .

4 Theory Model and Methods

We develop the model used for our analysis in three stages. First, Base models are defined. Second, we do the structural transform for both of the relative price spread model and the markup pricing model. Finally, we compare these models and choose the better model.

4.1 Base Model

George and King suggested an empirical specification----the markup pricing model (1):

$$M_t = a_0 + a_1 P_{rt} + a_2 IC_t + \epsilon_{1t} \quad (1)$$

And later Wohlgenant and Mullen proposed the new model ----the relative price spread model (2):

$$\frac{M_t}{P_{rt}} = b_0 + b_1 Q_t + b_2 \frac{IC_t}{P_{rt}} + \epsilon_{2t} \quad (2)$$

Where M_t is the marketing margin P_{ft}– P_{rt} at time t, Q_t is per capita quantity of pork produced, IC_t is an index of marketing costs and ϵ_{1t} , ϵ_{2t} are random errors.

Hall, Schmitz, and Cothern pointed out in 1979, that the retail firm concentration was positively correlated with the wholesale-retail marketing margins. Based on Capps, Byrnes, and Williams's conclusions, we further develop the models by including the four-firm concentration ratios for pork slaughter.

$$M_t = a_0 + a_1 P_{rt} + a_2 IC_t + a_3 (CR_4)_t + \epsilon_{3t} \quad (3)$$

$$\frac{M_t}{P_{rt}} = b_0 + b_1 Q_t + b_2 \frac{IC_t}{P_{rt}} + b_3 (CR_4)_t + \epsilon_{4t} \quad (4)$$

4.2 Box-Cox Transform

The usual Box-Cox method is to find the maximum likelihood power transformations of the dependent variable in a regression model.

When the dependent variable Y is known to be positive, the following transformation can be used:

$$y_i^{(\lambda)} = \begin{cases} \frac{(y_i^{\lambda} - 1)}{\lambda} & \text{when } \lambda \neq 0 \\ \log(y_i) & \text{when } \lambda = 0 \end{cases}$$

In this paper, we would like to find the power transformations for some or all of the predictors in a regression model which is called the Box-Tidwell method. In the SAS procedure, the Box-Tidwell is recognized as a special case of Box-Cox. So we still call it Box-Cox, but need to differ from the usual Box-Cox.

Box-Cox regression model:

$$y_i^{(\lambda)} = \beta_0 + \beta_1 x_{1i}^{(\lambda 2)} + \beta_2 x_{2i}^{(\lambda 2)} + \varepsilon_i$$

Theoretically, we can allow for the different transformations for every variable. In this paper, we assume that the dependent variable and independent variables have the same transformation.

4.3 Residual Analysis

4.3.1 Checking the Homoscedasticity Assumption

One of the assumptions necessary for the validity of regression inferences is that the error term ε has constant variance σ^2 for all levels of the independent variables. So the error term has to be homoscedastic. Many statistical tests for heteroscedasticity have been developed. Plots of the residuals are the usual method which can frequently reveal the presence of heteroscedasticity(Mendenhall). The points in the residuals versus the fitted values distributing randomly indicates homoscedasticity, otherwise the dependent variable need some kind of transformation.

4.3.2 Checking the Normality Assumption

Another assumption for all the inferential procedures with the regression model is that the random error term ε is normally distributed with mean 0 and variance σ^2, and the observations are independently distributed. So we construct a normal probability plot by residuals versus the expected values of the residuals to check the assumption of normality (Mendenhall). If the errors are normally distributed, then the residuals will approximately equal its expected value. A linear trend on the normal probability plot indicates that the normality assumption is nearly satisfied. Otherwise some kinds of transformations are needed for the model.

4.4 Detecting Residual Correlation

Regression models of time series may pose a special problem. As emphasized by Mendenhall (p.412), because the time series tend to follow economic trends and seasonal cycles, the value of a time series at time t is often indicative of its value at time (t+1), which means the value of a time series at time t is correlated with its value at time (t+1). This will lead to residual correlation and with that we can't apply the standard least squares inference-making tools and have confidence in their validity. For this problem, Wooldridge suggests the regression the model

$$\varepsilon_t = \rho_0 + \rho\varepsilon_{t-1} + \mu_t$$

Then do the t-test for H_0: $\rho = 0$ VS H_a: $\rho \neq 0$. If the t-test indicates that some serial correlation exists in the model, we should adjust the model with Proc Autoreg in SAS.

4.5 Adjusted R-Squared

Most regression packages will report, along with the R-square, a statistic called the adjusted R-squared. It can be represented in terms of R^2 (Wooldridge).

$$\text{Adjusted } R^2 = 1 - \frac{(1-R^2)(n-1)}{n-k-1}$$

Wooldridge discussed the method, using Adjusted R-Squared to choose between Nonnested Models. The model which has higher adjusted R^2 fits better. And also Wooldridge mentioned that there is an important limitation in using Adjusted-R Squared to choose between nonnested models: we cannot use it to choose between different functional forms for the dependent variable. Because either R^2 or Adjusted R^2 measure the explained proportion of the total variation in whatever dependent variable we are using in the model. Comparing the adjusted R^2 from regressions with these different forms of the dependent variables cannot tell us anything about which model fits better.

5 Empirical modeling

We first present the results of the OLS version of model (3) and Model (4) executed using our data---- U.S. annual time-series data covering from 1970 to 2008, a total of 39 observations. These results are reported in table 1. From table 1, P_r and CR_4 are highly significant at the 5 percent significance level and are positively correlated with the farm-retail price spread. However, IC and IC/P are not significant at that level.

Table 1. Economics Estimates of Model (3) and (4) of the Farm-Retail Price Spread for Pork, 1970-2008

		Explanatory Variables					Statistics
Model	Intercept	P	Q	IC	IC/P	CR_4	Adjusted R^2
M	33.669 (13.988)[a]	0.278 (0.056)		-0.183 (0.117)		0.685 (0.125)	0.4778
M/P	-0.0641 (0.113)		0.00592 (0.00188)		-0.230 (0.138)	0.00842 (0.599×10^{-3})	0.8475

[a] Standard error of the coefficient
Observations=39
[b] Variable definitions given in Table

The plot ordered residuals with ordered quantile of N (0, 1) are Figure A.3 and Figure A.5, which are the Quantile-Quantile plots. The Studentized Residuals with predicted dependent variable are shown in Figure A.4 and Figure A.6 to check the Normality and Homogeneity assumptions.

The line in the Quantile-Quantile plot for model (3) (Figure A.3) is almost straight, except for the right tail which shows a slight downward bias of the line. From Figure A.4, we see some points outside of the [-2, 2], but we still can't find any significant pattern in residuals with the predicted values. We can't really tell for sure whether the data need to be transformed. Therefore, the Box-Cox transformation is used for both independent variable and dependent variables to check bias on the functional form.

For model (4), it is very clear that the data need to be transformed. By the Quantile-Quantile plot (Figure A.5), the right tail of the curve is almost on a straight line, the slope of which is 1. But the left tail of the curve is far outside of the straight line (slope=1). Moreover, the points are crowded on the left and right-hand sides. First, we doubt that the biggest outlier makes the points at the left tail far outside of the straight line. So we check all of the residuals and find that the biggest residual is 1973. When we put the dummy variable in that year and run the regression, the normality assumption was improved. But there still exists problems for homoscedasticity assumption. The residual analysis results are shown in Figure 7 and Figure 8. So taking a Box-Cox transformation for both independent variable and dependent variables will help determine which transformation of the data is needed.

Table 2. Box-Cox transform for Model 3 and Model 4

		Explanatory Variables					Box -Cox	
Model	Intercept	P	Q	IC	IC/P	CR	Lambda[a]	Sigma[b]
M	2.789	0.461		- 0.162		0.461	0.354	0.317
	(7.374)[c]	(0.262)		(0.107)		(0.220)	(0.706)	(1.020)
M/P	-1.682		0.02065		-0.194	0.026	0.737	0.047
	(1.090)		(0.05320)		(0.146)	(0.062)	(0.539)	(0.015)
M/P[d]	-1.401		0.141		-0.094	0.156	0.230	0.039
	(0.809)		(0.145)		(0.082)	(0.146)	(0.248)	(0.004)

[a] Box-Cox parameter
[b] Standard deviation of the error
[c] Standard error of the coefficient
[d] Do the Box-Cox for independent variables only

The Box-Cox results for both models are reported in table 2. The value of lambda for transforming both the independent and dependent variables for model (3) is 0.354, with a standard error 0.706. The value of lambda for transforming both independent and dependent variables for model (4) is 0.737, with a standard error 0.539. The value of lambda of only transforming the independent variables for model (4) is 0.230, with a standard error 0.248. The two-side t-test is conducted for the null and alternative hypotheses ($H_{0:}$ Lambda=0 vs H_a=Lamda≠0) for all three models with thirty-four degrees of freedom. The critical value is 2.03 at the 95% confidence level. Compared with the critical value, both t-tests indicate failure to reject the null hypothesis. Base on the t-test, we take Lambda=0 for these three models, which means taking log transformations for independent variables and dependent variables.

Also the two-sided t-test is conducted for the null and alternative hypotheses (H_0: Lambda=1 VS H_a=Lamda≠1) for all three models. Compared with the critical value, both t-tests indicate failure to reject the null hypothesis in the first two models. But the null hypotheses is rejected in the third model. So far, we cannot tell whether we should transform model (3) and model (4).

$$log(M_t) = a_0 + a_1 log(P_{rt}) + a_2 log(IC_t) + a_3 log(CR_4)_t + \epsilon_{5t} \qquad (5)$$

$$log(\frac{M_t}{P_{rt}}) = b_0 + b_1 log(Q_t) + b_2 log(\frac{IC_t}{P_{rt}}) + b_3 log(CR_4)_t + \epsilon_{6t} \qquad (6)$$

$$\frac{M_t}{P_{rt}} = b_0 + b_1 log(Q_t) + b_2 log(\frac{IC_t}{P_{rt}}) + b_3 log(CR_4)_t + \epsilon_{7t} \qquad (7)$$

Table 3 represents the OLS model after log transform. From Table 3, log(IC) and $log(\frac{IC}{P_r})$ are still not significant at 5 percent significant level in all three models. However, $log(P_r)$, $log(CR_4)$, $log(Q)$ are all highly significant at the 5 percent significant level in all the three models.

By comparing the adjusted R^2 in model (7) with model (4), we prefer the Box-Cox transformation on the independent variable to no transformation. Because the dependent variables are not the same in models (3) and (5), and models (4) and (6), we cannot compare the models directly. So we rewrite the models (5) and (6) into models (5^*) and (6^*) and represent the result in table 4. The adjusted R^2 for model (6*) is 0.8657 which is greater than 0.8475 in model (4), which indicates that the Box-Cox transformation on both the independent and dependent variable improves model (4). But the adjusted R^2 for model (5*) is lower than the adjusted R^2 in model (3). So far, the Box-Cox transformation has not improved model (3).

Table 3. Economics Estimates of Model (5), (6) and (7) of the Farm-Retail Price Spread for Pork, 1970-2008

		Explanatory Variables					Statistics
Model	Intercept	log(P)	log(Q)	log(IC)	log(IC/P)	log(CR)	Adjusted R^2
log(M)	0.764 (0.761)[a]	0.601 (0.116)		-0.149 (0.106)		0.367 (0.065)	0.4836
log(M/P)	-5.848 (1.053)		0.677 (0.228)		-0.100 (0.125)	0.637 (0.048)	0.8364
M/P	-2.407 (0.563)		0.370 (0.122)		-0.073 (0.067)	0.371 (0.026)	0.8572

[a] Standard error of the coefficient
Observations=39
[b] Variable definitions given in Table

Table 4. Economics Estimates of Model (5*) and (6*) of the Farm-Retail Price Spread for Pork, 1970-2008

		Explanatory Variables		Statistics
Model	Intercept	$\exp(\widehat{log(M_t)})$	$\exp(\widehat{log(\frac{M_t}{P_{rt}})})$	Adjusted R^2
M	1.1096 (23.554)	0.9913 (0.2470)		0.2845
M/P	0.0047 (0.0369)		0.9943 (0.0634)	0.8657

The residuals were analyzed to check the Homoscedasticity and Normality Assumptions for models (5), (6) and (7). The results are presented in Figures A.9-A.14. For model (5), all points are almost on the slope=1 line in the Quantile-Quantile plot and distributed randomly in Figure A.10. There is no pattern in the residuals VS time plot which is shown in Figure A.15. The only problem is there are some outliers that lie on the line [-2, 2]. For models (6) and (7), the left tail of the Quantile-Quantile plot is downward biased relative to the line. Also the residual-predicted values are distributed along the break line. As shown in Figures A.15 and A.16, the points in the residuals VS time plot are distributed according to some pattern, which indicates that a trend exists in the model.

As indicated by Wooldridge (Chapter 10), unobserved, trending factors that affect the dependent variable in the regression may also be correlated with one or more explanatory variables on the right-hand side of the regression. Spurious regression results may therefore occur if variables are not detrended. Including a trend variable in the model eliminates the problems of spurious correlation.

So we further modified the model by including the linear time trend into model (6) and (7).

$$log(\frac{M_t}{P_{rt}}) = b_0 + b_1 log(Q_t) + b_2 log(\frac{IC_t}{P_{rt}}) + b_3 log(CR_4)_t + b_4 t + \epsilon_{8t} \qquad (8)$$

$$\frac{M_t}{P_{rt}} = b_0 + b_1 log(Q_t) + b_2 log(\frac{IC_t}{P_{rt}}) + b_3 log(CR_4)_t + b_4 t + \epsilon_{9t} \qquad (9)$$

The OLS results for model (8) and (9) are presented in Table 5. From Table 5, $log(CR_4)$ is not significant at 5 percent significant level in the model (8) and (9), but it is significant in models (6) and (7). The $log(Q)$ variable is still highly significant at the 5 percent significant level in these two models. The $log(\frac{IC_t}{P_{rt}})$ variable turns to be highly significant with the same sign as in model (6) and (7). Time trend is highly significant. The adjusted R^2 of model (8) and (9) are 0.9031 and 0.9042.

Table 5. Economics Estimates of Model (8) and (9) of the Farm-Retail Price Spread for Pork, 1970-2008

		Explanatory Variables				Statistics
Model	Intercept	log(Q)	log(IC/P)	log(CR_4)	t	Adjusted R^2
log(M/P)	-5.807 (0.811)[a]	1.041 (0.190)	-0.299 (0.104)	0.100 (0.113)	0.012 (0.0029)	0.9031
M/P	-2.387 (0.461)	0.546 (0.108)	-0.169 (0.059)	0.112 (0.064)	0.0070 (0.0016)	0.9042

[a] Standard error of the coefficient
Observations=39
[b] Variable definitions given in Table

Analyzing the residuals of models (8) and (9) (Figure A.18-20), combined with the analysis results of model (6), indicates there are several outliers in our data.

The years 1998 and 1999 were significant for the swine industry because of extraordinary increases in supply of hogs that constrained slaughter capacity causing significantly higher slaughter costs and therefore price spreads for pork.

In other studies, 1973 has also been found to be an outlier, perhaps because of shocks to livestock industry in that time period.

So we put two dummy variables into the model, d_1 and d_2. The d_1 is equal to 1, for in 1998 and 1999, and otherwise equal to 0. d_2 is equal to 1, when year equals 1973, and equal to 0 otherwise.

$$log(M_t) = a_0 + a_1 log(P_{rt}) + a_2 log(IC_t) + a_3 log(CR_4)_t + a_4 d_1 + a_5 d_2 + \epsilon_{10t} \quad (10)$$

$$log(\frac{M_t}{P_{rt}}) = b_0 + b_1 log(Q_t) + b_2 log(\frac{IC_t}{P_{rt}}) + b_3 log(CR_4)_t + b_4 t + b_5 d_1 + b_6 d_2 + \epsilon_{11t} \quad (11)$$

$$\frac{M_t}{P_{rt}} = b_0 + b_1 log(Q_t) + b_2 log(\frac{IC_t}{P_{rt}}) + b_3 log(CR_4)_t + b_4 t + b_5 d_1 + b_6 d_2 + \epsilon_{12t} \quad (12)$$

The OLS results are represented in Table 6. The dummy variables are both highly significant. d_1 is positively correlated with the margins and d_2 is negatively correlated with the margins. The other variables are all significant except $log(CR_4)$.

Table 6. Economics Estimates of Model (10), (11) and (12) of the Farm-Retail Price Spread for Pork, 1970-2008

Independent Variables	Models log(M)	log(M/P)	M/P
Intercept	0.258	-5.626	-2.172
	(0.711)	(0.675)	(0.405)
log(P)	0.681		
	(0.115)		
log(Q)		0.968	0.491
		(0.154)	(0.093)
log(IC)	-0.136		
	(0.099)		
log(IC/P)		-0.324	-0.162
		(0.089)	(0.053)
log(CR_4)	0.376	0.139	0.121
	(0.059)	(0.091)	(0.055)
t		0.013	0.006
		(0.002)	(0.001)
d_1	0.103	0.081	0.061
	(0.046)	(0.037)	(0.022)
d_2	-0.156	-0.206	-0.086
	(0.068)	(0.052)	(0.031)

Table 6. (Continued)

Statistics	Models		
	log(M)	log(M/P)	M/P
Adjusted R^2	0.5880	0.9390	0.9329

[a] Standard error of the coefficient
Observations=39
[b] Variable definitions given in Table

The residual analysis results for these models are reported in Figure A.22-27. All the points in these models are almost on the straight line, which means the homoscedasticity assumptions are almost valid. And the Normality Assumptions are also almost valid though the figure still indicates some outliers.

Serial correlation in error term of multiple regression models can produce problems, such as OLS no longer produces BLUE estimates and standard errors of OLS estimates are biased and inconsistent. So a test for serial Correlation is very necessary.

We take the t-test to test the truth of $H_0: \rho = 0$ against $H_1: \rho \neq 0$ for all three models. The results of the tests are reported in table 7. The p-value of all three models are greater than α=0.05, which is not significant at 5 percent significant level. So there is not enough evidence to say that serial correlation exists in these models. So we can report models (10), (11), (12) as our final models.

Table 7. Serial Correlation Test for Model (10), (11) and (12)

	ρ	D-W	p-value
Model (10)	0.072 (0.165)	1.840	0.6647
Model (11)	0.066 (0.165)	1.856	0.6952
Model (12)	0.150 (0.164)	1.693	0.3662

6 Results

Model (10), (11), (12) were reported to analyze the determinants of margins.

From table 6, not all variables are significant at the 5 percent significance level. In particular, the log of quantity, log of industry costs deflated by retail price, time trend and two dummy variables are highly significant at 5% percent significant level. Only the log of four-firm concentration ratio is not significantly related to the retail-farm margins.

In order to compare these models, some basic statistical transforms are performed on model (11).

We rewrite model (11) in two ways. First, we rewrite it by shifting the log (P_{rt}) from left side to the right side of the equation. So the new model can write as:

$$log(M_t) = (b_0 + log(P_{rt})) + b_1 log(Q_t) + b_2 log(\frac{IC_t}{P_{rt}}) + b_3 log(CR_4)_t + b_4 t + b_5 d_1 + b_6 d_2 + \epsilon_{11t} \quad (13)$$

Model (13) has the same dependent variables as model (10). So we can compare model (10) with model (13) by adjusted R^2 directly. The adjusted R^2 for model (10) is 0.5880, which means that 58.80% of variation in the dependent variable in the model is explained by the variables which were used for the regression. The adjusted R^2 for model (13) is higher, equal to 0.9390. So the set of independent variables in model (13) cannot explain only 6.10% of the variation of $\log(M_t)$. So model (13) is preferred comparing to model (10).

Rewrite model (11) by obtaining the fitted value of the regression of $log(\frac{M_t}{P_{rt}})$ on the independent variables in model (11) and then run model (14):

$$\frac{M_t}{P_{rt}} = \beta_0 + \beta_1 \exp(\widehat{log(\frac{M_t}{P_{rt}})}) + \varphi_t \qquad (14)$$

Model (14) is another form of model (11) with a different form of the dependent variable. The results of the OLS of model (14) are displayed in Table 8. Now model (12) and model (14) have the same dependent variables with different numbers of independent variables. We can compare these two models by Adjusted R^2 directly. The adjusted R^2 of model (12) is reasonably high 0.9329. And the adjusted R^2 of model (14) is 0.9519 which is 0.019 higher than the adjusted R^2 of model (12). By comparing models (12) and (14), Model (14) is preferred.

All considered, Model (11) is better than the other two models (10) and (12) to specify the retail-farm price spread.

Table 8. Economics Estimates of Model (14) of the Farm-Retail Price Spread for Pork, 1970-2008

		Explanatory Variables	Statistics
Model	Intercept	$\exp(\widehat{log(\frac{M_t}{P_{rt}})})$	Adjusted R^2
M/P	-0.006 (0.022)	1.012 (0.037)	0.9519

7 Conclusion

The purpose of the study is to find a model which can describe the change of the farm-retail price spread of pork most accurately.

Based on the basis model from George and King, and Wohlgenant and Mullen, we extended the model by including the four-firm concentration, time trend variables and importing two dummy variables. Meantime, we test the structural change of the model.

By statistical analysis, the log transform of the relative pricing model fits best. It can explain how the price spread changes when the demand, supply and marketing costs change simultaneously. The result is consistent with the analysis of Gardner(1975) and Wohlgenant and Mullen (1987).

The findings from the descriptive statistics, residual analysis, and regression results indicate that the log of farm-retail price spread is positively correlated to increases in log of retail price and positively related to log of quantity of the farm input. And the relationship between the price spread and industry costs is indeterminate. This result confirms the findings of Wohlgenant (2007) that the sign of the effect of a change in marketing input costs on price spread is indeterminate. It depends upon whether substitution effects are larger (smaller) than output effects.

Changes in packer concentration had a positive effect on farm-retail price spread, though not statistically significant. The problem of non-significance may be due to use of a concentration variable for pork slaughter, but not retail concentration. As we know, the pork slaughter concentration ratio would be hypothesized to have a positive relation with farm-wholesale margins, but it would not directly reflect the change of the farm-retail price spread. Additional study and more detailed data may be necessary to fully answer this question.

Most significantly, the results point to the strong possibility of spurious correlation between the price spread and concentration variables. When the time trend is excluded, the concentration variable has a positive, highly significant effect on the price spread. However, when the variables are detrended by including a linear time trend on the model, the concentration ratio's significance is dramatically reduced and turns out to be statistically insignificant in the preferred model. This result strongly suggests other possibly unobserved variables, such as exports and technical change, correlated with trend are spuriously indicating that concentration ratio has a significant effect on the price spread. Previous studies (e.g. Capps, Byrnes, and Williams) incorporating concentration ratio variables have failed to detrend their variables, possibly producing spurious results. A major implication of this study is that variables used in the regression need to be detrended in estimation.

References

Armah, S. (2007), "An Empirical Analysis of Recent Changes in US Beef Marketing Margins", Unpublished paper, University of Illinois at Urbana-Champaign

Brester, G. W., and J.M. Marsh (1983), "A Statistical Model of the Primary and Derived Market Levels in the U.S. Beef Industry." *Western Journal of Agricultural Economics* 8:34-49.

Capps, O., P. Byrnes, and G. Williams (1995): "Analysis of Marketing Margins in the U.S. Lamb Industry." *Agricultural and Resource Economics Review* 24(2), 232-240.

Elizak, H. (1996), "Food cost review", *Agricultural Economics Report* No. 729 (U.S. Department of Agriculture, Washington, D.C.)

Gardner, B. L. (1975), "The farm-retail price spread in a competitive food industry." *American Journal of Agricultural Economics* 57:399-409.

George, P. S., and G. A. King (1971), *Consumer Demand for Food Commodities in the United States with Projections for 1980*, Giannini Foundation Monograph No.26, University of California, Berkeley.

Hahn, W. (2004): "Beef and Pork Values and Price Spreads Explained." *Electronic Outlook report of the Economic Research Service*, UCDA, LDP-M(118).

Hall, L., A. Schmitz, and J. Cothern (1979): "Beef Wholesale-Retail Marketing Margin and Concentration." *Economica* 46, 295-300.

Heien, D. M. (1977), “Price Determination Processes for Agricultural Sector Models”, *American Journal of Agricultural Economics* 59:126-132.

Johnson, R. and G. S. Backer(2009), *Livestock Marketing and Competition Issues*, Congressional Research Service.

Mendenhall, W., and T. Sincich (2003), “Residual Analysis”, in: A Second Course in Statistics Regression Analysis, sixth edition. Chapter 8

Tomek, W. G., and K. L. Robinson (1990), “Marketing margins for farm products”, in: Agricultural Product Prices, third edition. (Cornell Univ. Press, Ithaca, N.Y.) Chapter 6

Waugh, F. V. (1964), “Demand and Price Analysis”, *Technical Bulletin* NO. 1316. (U.S. Department of Agriculture, Washington, D.C.)

Wohlgenant, M. K., and J. D. Mullen (1987), “Modeling the farm-retail price spread for beef”, *Western Journal of Agricultural Economics* 12:119-125

Wohlgenant, M. K. (2001), “Marketing Margins: Empirical Analysis”, *Bruce Gardner and Gordon Rausser, ed Handbook of Agricultural Economics* 1:934-970.

Wohlgenant, M. K. (2007), “Understanding Causes of Higher Price Spreads for Food”, Unpublished paper, North Carolina State University

Wohlgenant, M. K. (2010), “Modeling the Effects of Restricting Packer-Owned Livestock in the U.S. Swine Industry”, *American Journal of Agricultural Economics* 1–10; doi: 10.1093/ajae/aap035

Wooldridge, J. M. (2009), “Multiple Regression Analysis: Further Issues”, in: Introductory Econometrics, A Model Approach, fourth edition. Chapter 6

Wooldridge, J. M. (2009), “Basic Regression Analysis with Time Series Data”, in: Introductory Econometrics, A Model Approach, fourth edition. Chapter 10

U.S.Department of Labor. *Employment, Hours, and Earnings from the Current Employment Statistics survey (National)*. Washington DC, various issues.

U.S.Department of Labor. *Consumer Price Index - All Urban Consumers.* Washington DC, various issues.

U.S.Department of Labor. *Producer Price Index-Commodities.* Washington DC, various issues.

U.S. Department of Agriculture. *Packers and Stockyards Statistical Reports.* Washington DC, various issues.

U.S. Department of Agriculture. *Meat Price Spreads Data*. Washington DC, various issues.

U.S. Department of Agriculture. *Red Meats: Supply and Utilization*. Washington DC, various issues.

U.S. Department of Agriculture. (1975), *Farm-retail spreads for food products.* Agric. Marketing Serv. Miscell. Publ. 741.

Appendix

Table A1: Variable Definition and Symbols

Symbol	Variable Description
M	Farm-Retail Margin, c/lb Deflated
log(M)	The log transform of M
M_P_r	Farm-Retail Margin / Retail price
$\log(M/P_r)$	The log transform of M / Retail price
P_f	Farm price, c/lb Deflated
P_r	Retail price, c/lb Deflated
$\log(P_r)$	The log transform of P_r
Q	Per capita quantity of pork disappearance
log(Q)	The log transform of Q
IC	Index of marketing costs, Deflated, $
log(IC)	The log transform of IC
IC/P	Index of marketing costs / Retail price
log(IC/P)	The log transform of IC/P
CR_4	Top 4 concentration ration for Pork, %
$\log(CR_4)$	The log transform of CR_4
t	Time Trend
d_1	If year=1998 or 1999, d_1=1; else d_1=0
d_2	If year=1973, d_2=1; else d_2=0

List of raw data for:(1) Retail price of US pork production, (2) Retail price of US pork production deflated by the Consumer Price Index in US, then multiple 100, (3) Net farm value of US pork production, (4) Farm-retail price spread, (5) Farm-retail price spread deflated by Consumer Price Index in US then multiple 100, (6) Earnings of employees in animal slaughtering and processing (7) Index of earnings of employees in animal slaughtering and processing, (8) Producer price index of fuels and related products and power, (9) Industry Cost of US pork (10) Industry Cost of US pork deflated by the Consumer Price Index in US, then multiple 100, (11) Per capita quantity of US pork disappearance, (12)US Consumer Price Index divided by 100.

We estimate the relationship between the average wage of production workers in animal slaughter and processing and the average wage of production workers in nondurable manufacturing industries which based on observations from 1976-1982 only. Because when plot the two series, they move very closely together for that time period but start to deviate significantly from one another after 1982. After that we used the intercept and slope estimates from the output then to back-fill values for wages for animal slaughter and processing to 1970. The assumption that seems reasonable is to assume the relationship from 1976-1982 extends backward to 1970.

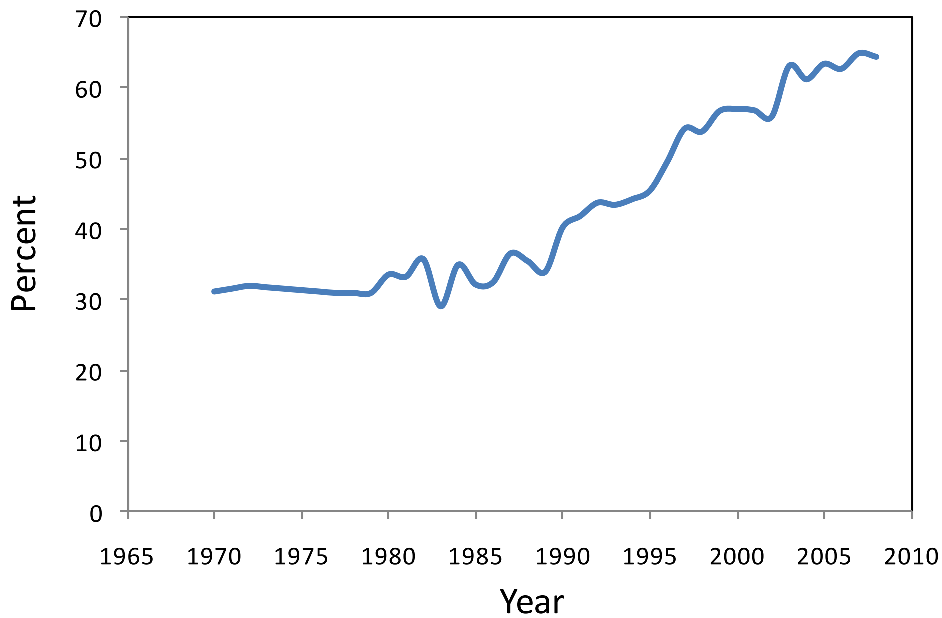
The data for CR_4 is only available from 1980 to 2007. And also we have the data in 1967, 1972, 1977. So we get the lost data, by assuming the CR_4 increase or decrease linearly from 1967 to 1980. And the 2008 data was estimated from the CR_4 graph in GIPSA's report in 2008.

Year	Retail price	Retail price/CPI*	Net farm value	Total Price Spread
1970	77.38333333	199.2276336	39.275	38.10833333
1971	69.81666667	172.4578016	32.01666667	37.8
1972	82.66666667	197.7277257	46.3	36.40833333
1973	109.2083333	245.826299	67.65833333	41.55
1974	107.8166667	218.6211558	57.2	50.60833333
1975	134.6333333	250.1315993	79.85833333	54.775
1976	134.0416667	235.4361827	70.95	63.09166667
1977	125.35	206.7913115	65.575	59.775
1978	143.55	220.0281007	76.55833333	66.99166667
1979	152.4754333	210.0694834	78.40766667	74.06776667
1980	147.51165	179.0552094	74.2825898	73.2290602
1981	161.2392	177.3158358	82.64190686	78.59729314
1982	185.5555667	192.2191644	103.4478918	82.1076749
1983	179.6660333	180.4177741	89.78844892	89.87758441
1984	171.4400833	164.9519724	90.95187451	80.48820882
1985	170.8172333	158.7520756	83.85006967	86.96716367
1986	188.7560167	172.0787207	96.69611487	92.0599018
1987	199.35365	175.4616253	97.11296322	102.2406868
1988	194.0460167	164.0634256	81.50710463	112.538912
1989	193.4553	156.0857662	82.68581567	110.7694843
1990	224.9131667	172.1384017	102.3115693	122.6015973
1991	224.1725667	164.6310159	91.97561042	132.1969563
1992	209.4928167	149.3088911	79.52980525	129.9630114
1993	209.0784333	144.7159947	85.12410065	123.9543327
1994	209.5369	141.3640749	73.87870488	135.6581951
1995	206.0895833	135.2441759	78.33195782	127.7576255
1996	233.7122	148.995718	99.39053428	134.3216657
1997	244.9710833	152.606188	95.33556507	149.6355183
1998	242.6932833	148.8839732	61.17036153	181.5229218
1999	241.4423558	144.9378824	60.37357798	181.0687779
2000	258.2004867	149.9494672	79.49838247	178.7021042
2001	269.3949175	152.164698	81.19912386	188.1957936
2002	265.7516552	147.7492523	61.93263837	203.8190168
2003	265.7955694	144.4541138	69.54275655	196.2528128
2004	279.2	147.7965504	92.1	187.1
2005	282.7	144.7578408	88.0	194.7
2006	280.8	139.2972303	83.3	197.5
2007	287.1	138.4626514	82.0	205.1
2008	293.7	136.4625291	82.5	211.2

Year	Price Spread/CPI*	Wage price	Wage price index	Fuel price index
1970	98.11199313	3.614480382	48.84432948	15.3
1971	93.37175793	3.774330328	51.0044639	16.6
1972	87.08391469	3.952734287	53.4153282	17.1
1973	93.52841868	4.142556098	55.98048782	19.4
1974	102.6191281	4.413016499	59.6353581	30.1
1975	101.7649791	4.714875997	63.7145405	35.4
1976	110.8167447	4.91	66.35135135	38.3
1977	98.61149299	5.32	71.89189189	43.6
1978	102.6823349	5.72	77.2972973	46.5
1979	102.0451435	6.17	83.37837838	58.9
1980	88.88819769	6.76	91.35135135	82.8
1981	86.43397339	7.21	97.43243243	100.2
1982	85.05629306	7.4	100	100
1983	90.25364125	7.26	98.10810811	95.9
1984	77.44215089	7.09	95.81081081	94.8
1985	80.82450155	7.09	95.81081081	91.4
1986	83.92606713	7.18	97.02702703	69.8
1987	89.98740219	7.31	98.78378378	70.2
1988	95.15021098	7.45	100.6756757	66.7
1989	89.37227271	7.64	103.2432432	72.9
1990	93.83373736	7.87	106.3513514	82.3
1991	97.08466799	8.03	108.5135135	81.2
1992	92.62672311	8.24	111.3513514	80.4
1993	85.79638878	8.41	113.6486486	80
1994	91.52180477	8.62	116.4864865	77.8
1995	83.83963175	8.88	120	78
1996	85.63247031	9.04	122.1621622	85.8
1997	93.21633282	9.26	125.1351351	86.1
1998	111.3580626	9.56	129.1891892	75.3
1999	108.6956145	9.88	133.5135135	80.5
2000	103.7809249	10.27	138.7837838	103.5
2001	106.3002835	10.53	142.2972973	105.3
2002	113.3167254	10.91	147.4324324	93.2
2003	106.6591374	11.3	152.7027027	112.9
2004	99.0427456	11.53	155.8108108	126.9
2005	99.69703435	11.47	155	156.4
2006	97.97436957	11.49	155.2702703	166.7
2007	98.9156733	11.81	159.5945946	177.6
2008	98.13035803	12.34	166.7567568	214.6

Year	IC	IC/CPI*	Quantity	CPI*	CR4
1970	35.426598	91.2077177	72.399367	0.3884167	31.2
1971	37.242678	91.9950885	79.058692	0.4048333	31.6
1972	38.889197	93.017812	71.271015	0.4180833	32
1973	41.348293	93.0743786	63.47653	0.44425	31.8
1974	47.821215	96.9676543	68.553897	0.4931667	31.6
1975	52.388724	97.3315825	55.713835	0.53825	31.4
1976	55.130811	96.8339768	58.526358	0.5693333	31.2
1977	60.575135	99.931485	60.467172	0.6061667	31
1978	64.978378	99.5964415	60.213846	0.6524167	31
1979	73.587027	101.382816	68.684404	0.7258333	31
1980	87.930811	106.733738	73.263973	0.8238333	33.6
1981	98.539459	108.364508	69.829018	0.9093333	33.3
1982	100	103.59116	62.569874	0.9653333	35.8
1983	97.224865	97.6316635	65.951143	0.9958333	29.1
1984	95.406486	91.7958497	65.507646	1.0393333	35
1985	94.046486	87.4037978	65.977177	1.076	32.2
1986	86.136216	78.5257612	62.343562	1.0969167	32.5
1987	87.35027	76.881564	62.7032	1.1361667	36.6
1988	87.085405	73.6295966	67.025348	1.18275	35.5
1989	91.105946	73.507117	66.394439	1.2394167	34
1990	96.730811	74.0334033	63.615747	1.3065833	40.3
1991	97.588108	71.6681332	64.147758	1.3616667	41.9
1992	98.970811	70.5380846	67.409126	1.4030833	43.8
1993	100.18919	69.3470768	66.492479	1.44475	43.5
1994	101.01189	68.1476754	67.181403	1.48225	44.3
1995	103.2	67.7239418	66.337781	1.5238333	45.5
1996	107.6173	68.6079566	62.028353	1.5685833	49.6
1997	109.52108	68.2268065	61.364103	1.60525	54.3
1998	107.63351	66.0294546	66.084784	1.6300833	53.9
1999	112.30811	67.4185742	67.669668	1.6658333	56.8
2000	124.67027	72.4020347	65.524019	1.7219167	57.1
2001	127.49838	72.0160292	64.413339	1.7704167	56.9
2002	125.73946	69.9070382	66.128987	1.7986667	56
2003	136.78162	74.3378378	66.508685	1.84	63.2
2004	144.24649	76.3579266	65.721649	1.8890833	61.3
2005	155.56	79.6552166	64.43299	1.9529167	63.5
2006	159.84216	79.2933421	63.14433	2.0158333	62.8
2007	166.79676	80.4427767	65.721649	2.0734833	65
2008	185.89405	86.3723962	68.298969	2.1522392	64.5(estimate)

Figure A.1 Four-Firm Concentration Ratios:Pork sectors, 1970-2008



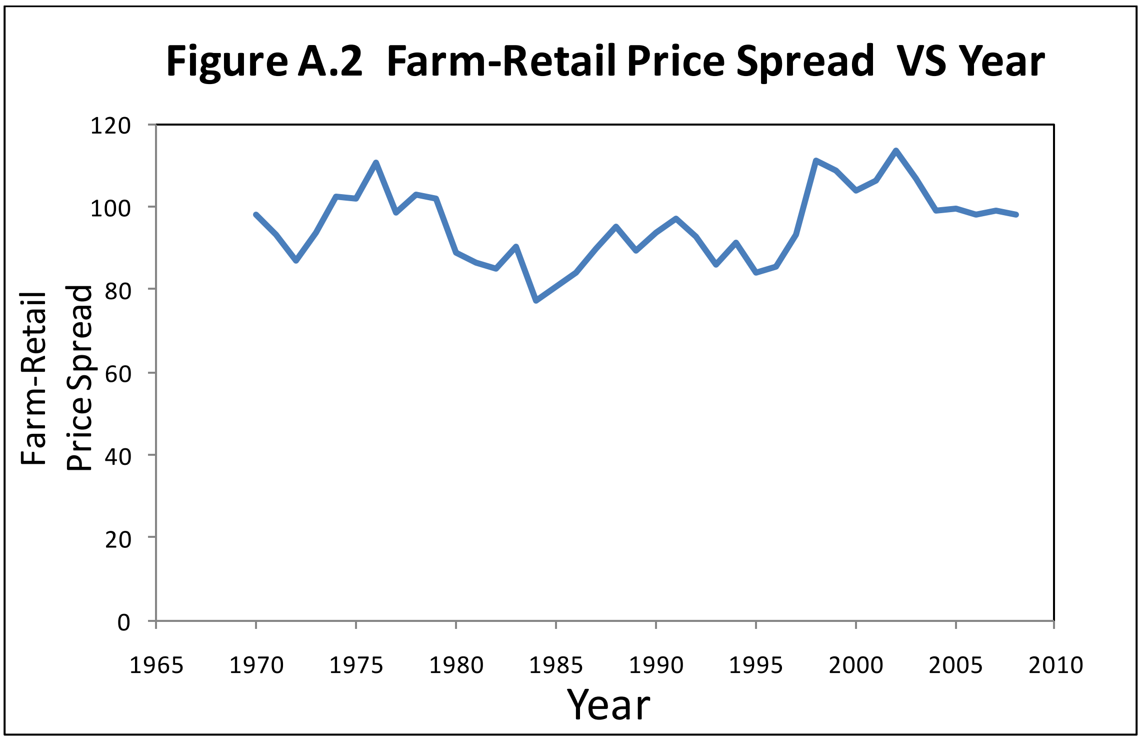


Figure A.2 Farm-Retail Price Spread VS Year
Farm-Retail Price Spread
120
100
80
60
40
20
0
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
Year

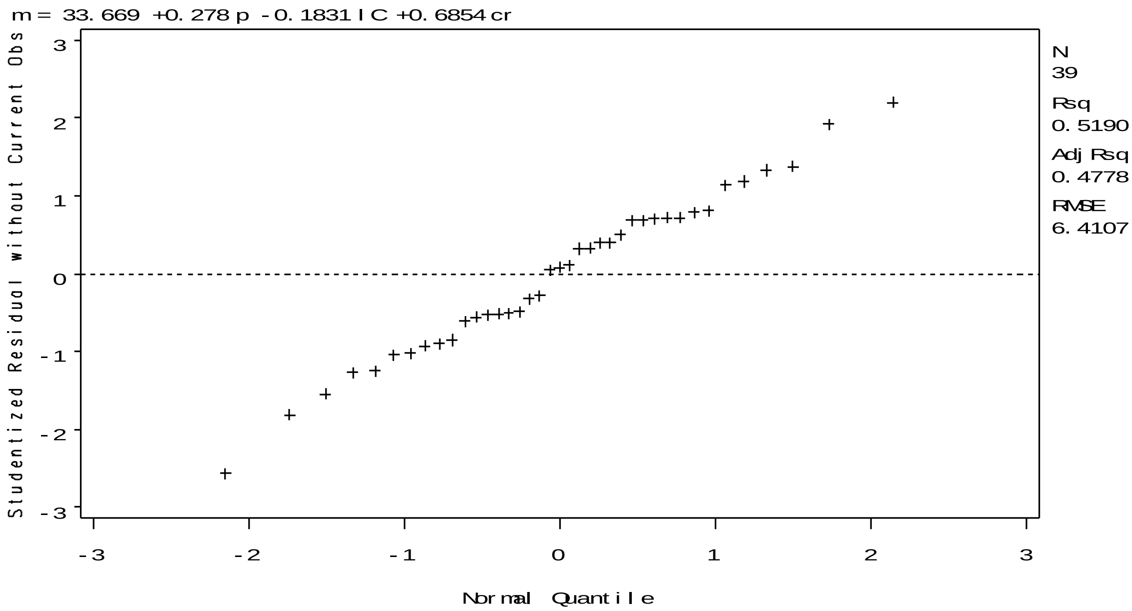


FIGURE A.3 **Quantile-Quantile plot for Model (3)**

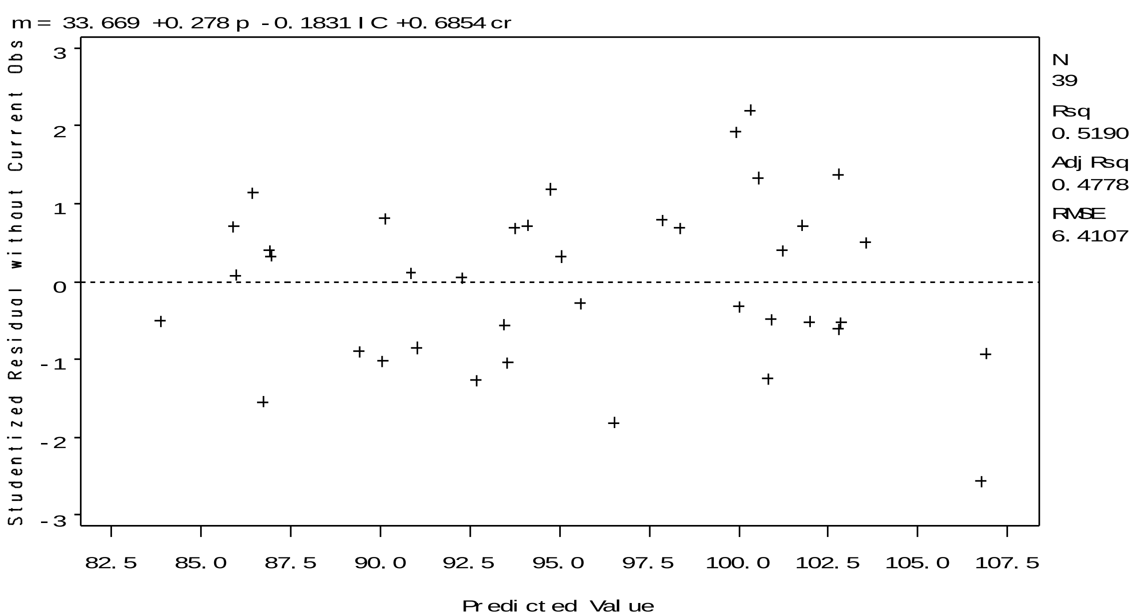


FIGURE A.4 **Studentized Residual VS predicted dependent variable for Model (3)**

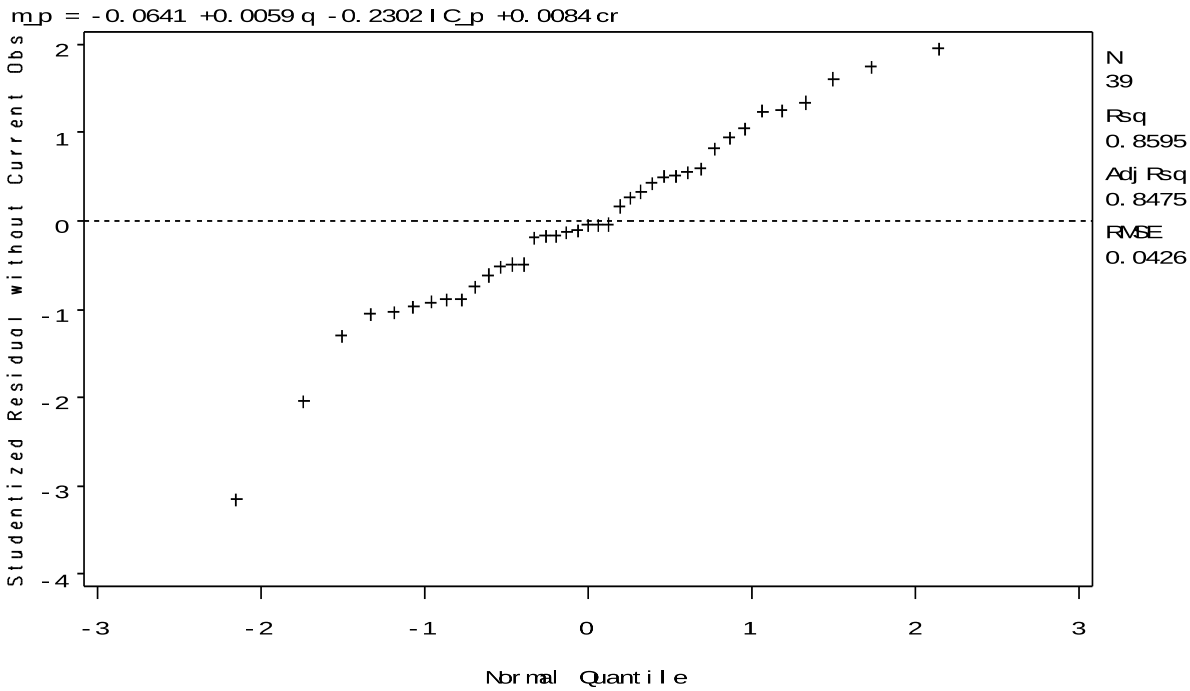


FIGURE A.5 **Quantile-Quantile plot for Model (4)**

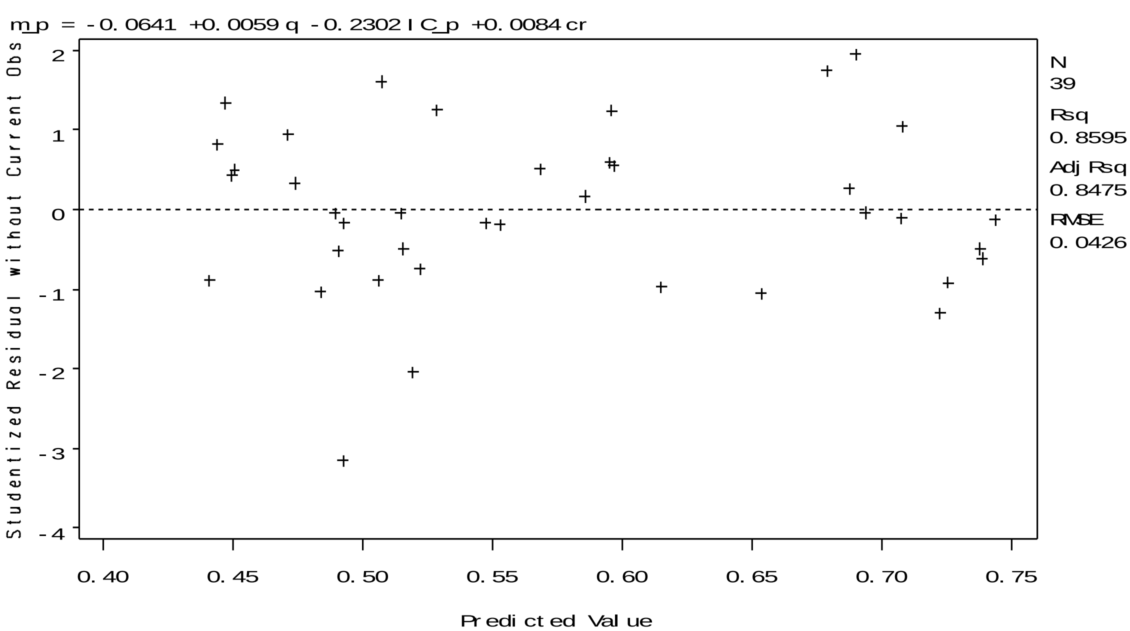
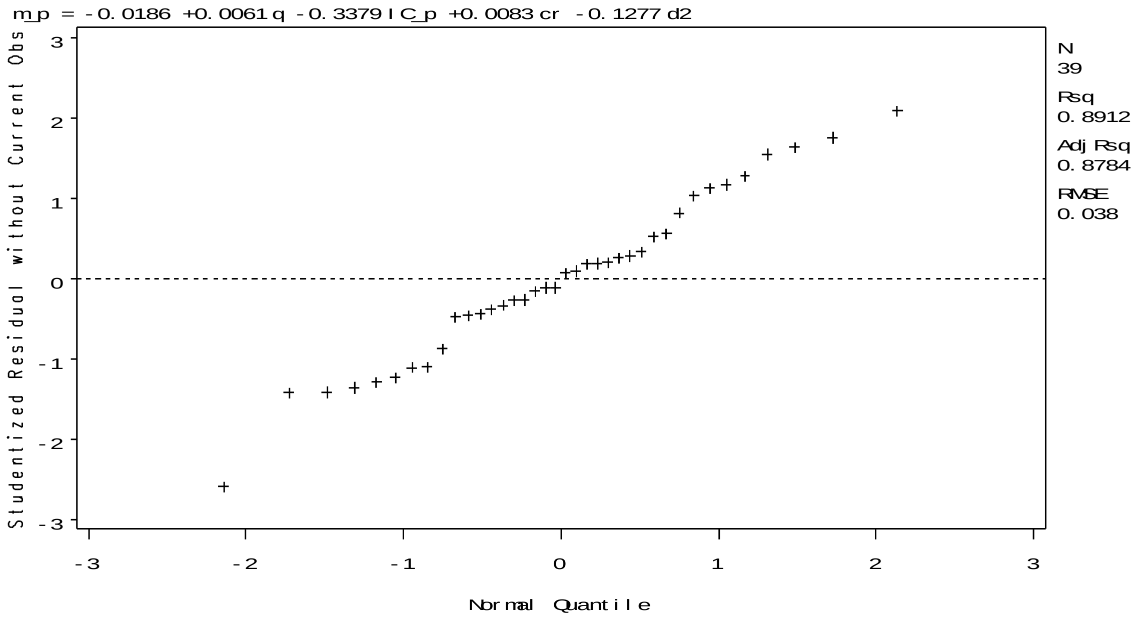
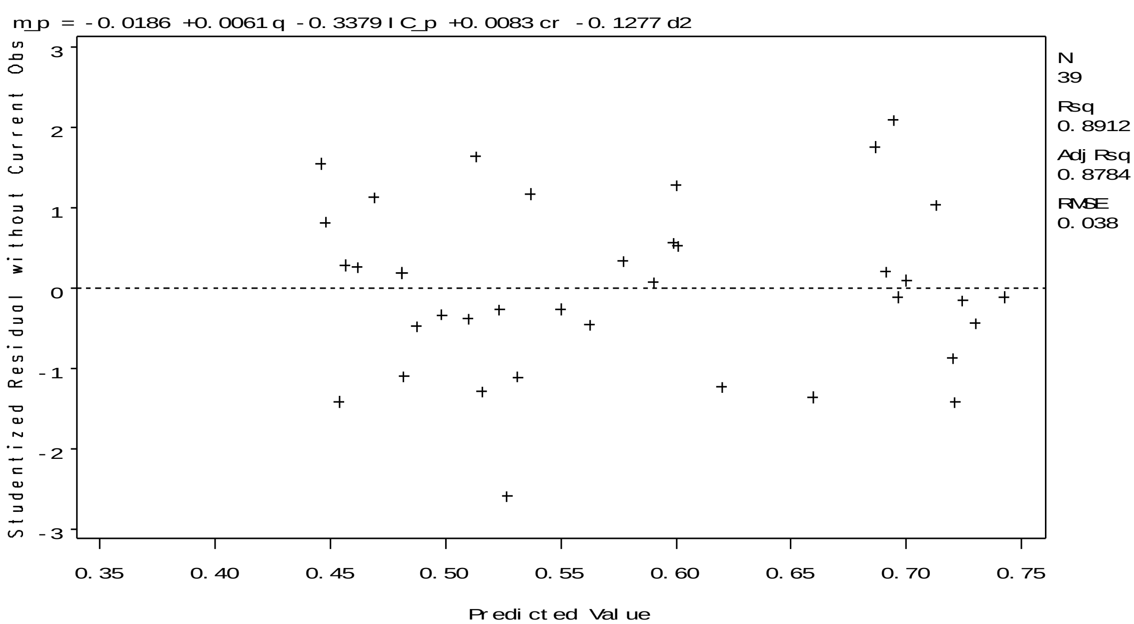


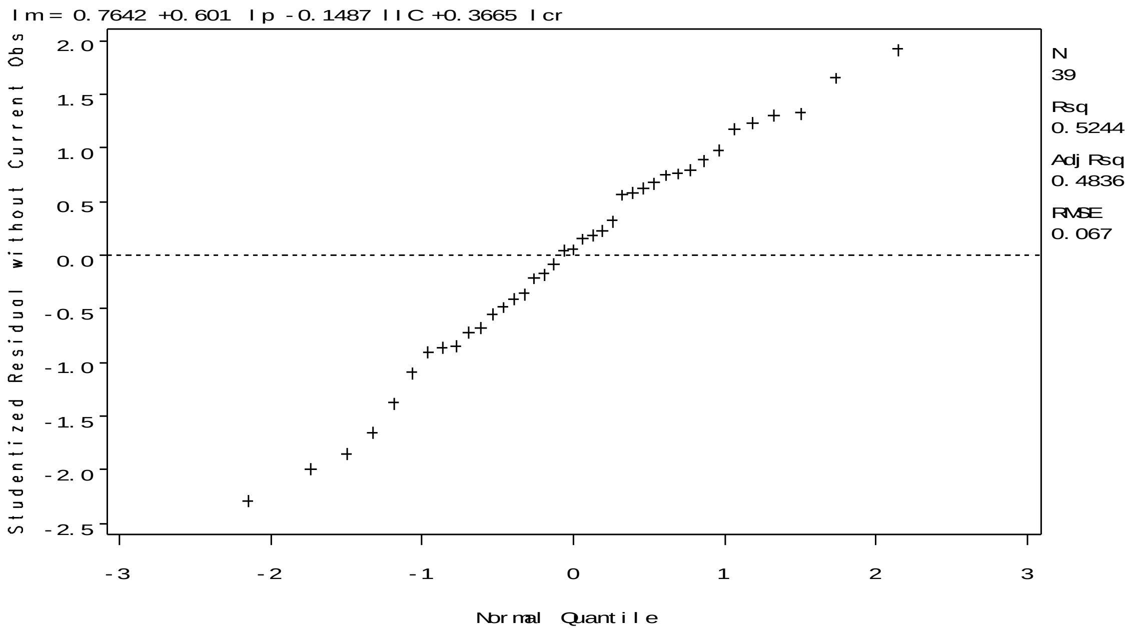
FIGURE A.6 **Studentized Residual VS predicted dependent Variable for Model (4)**



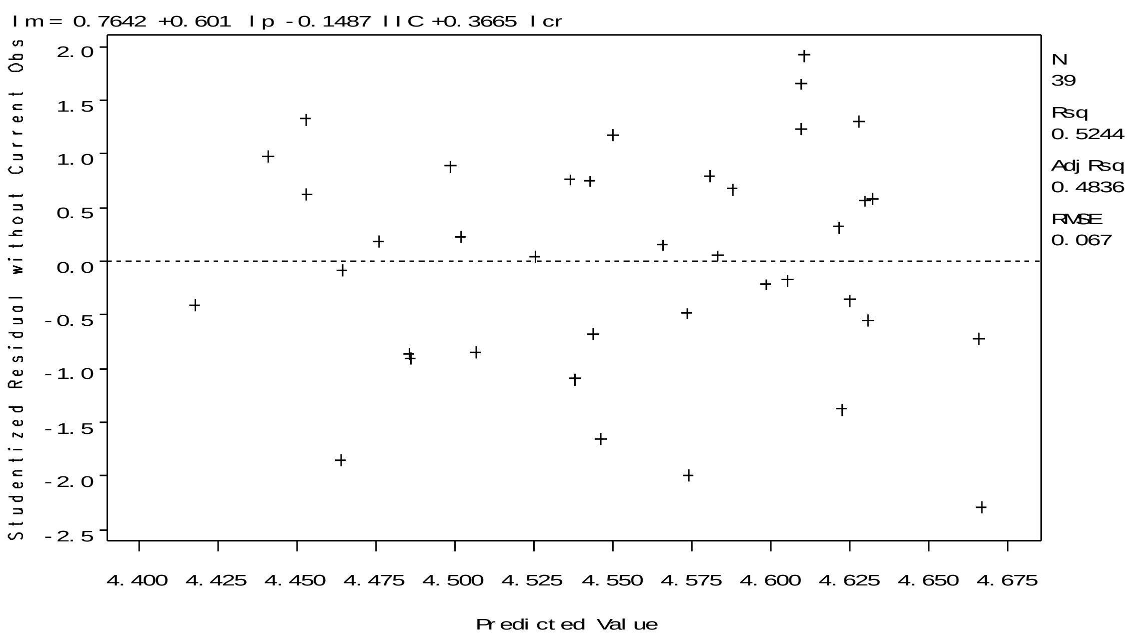
<u>FIGURE A.7</u> **Quantile-Quantile plot for Model (4) with putting 1973 as the dummy variable**



<u>FIGURE A.8</u> **Studentized Residual VS predicted dependent Variable for Model (4) with putting 1973 as the dummy variable**



<u>FIGURE A.9</u> **Quantile-Quantile plot for Model (5)**



<u>FIGURE A.10</u> **Studentized Residual VS predicted dependent Variable for Model (5)**

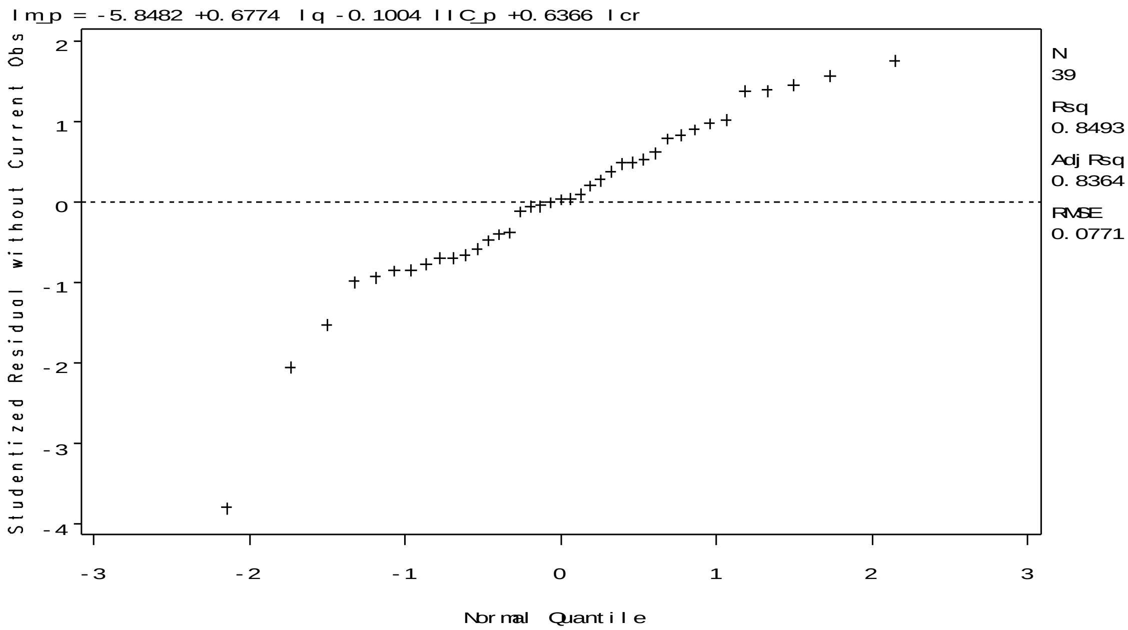


FIGURE A.11 Quantile-Quantile plot for Model (6)

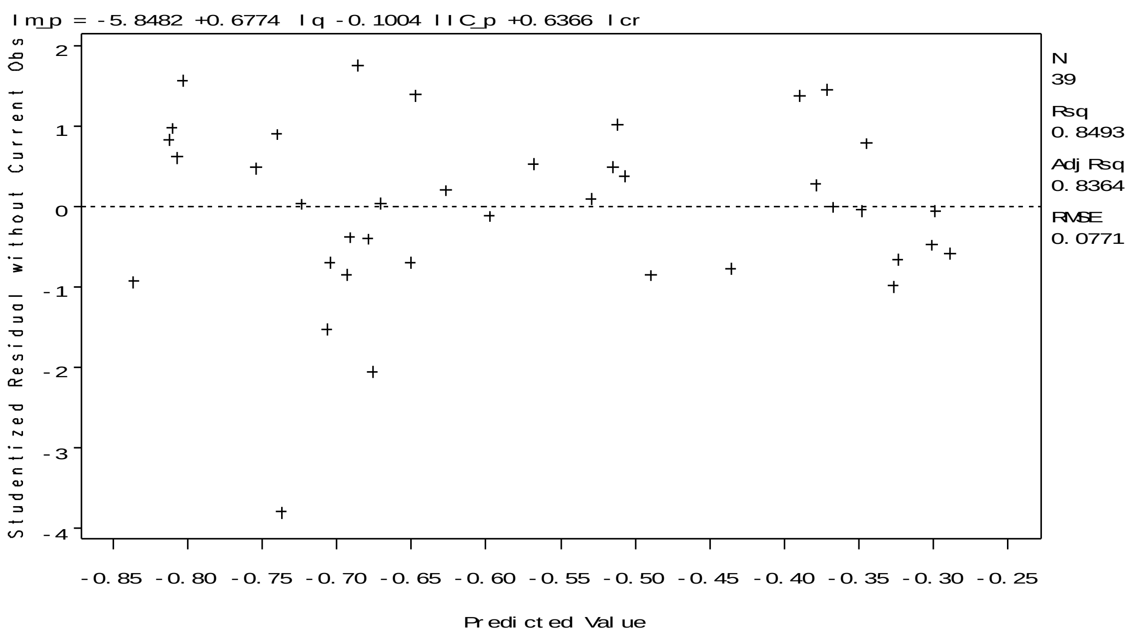


FIGURE A.12 Studentized Residual VS predicted dependent Variable for Model (6)

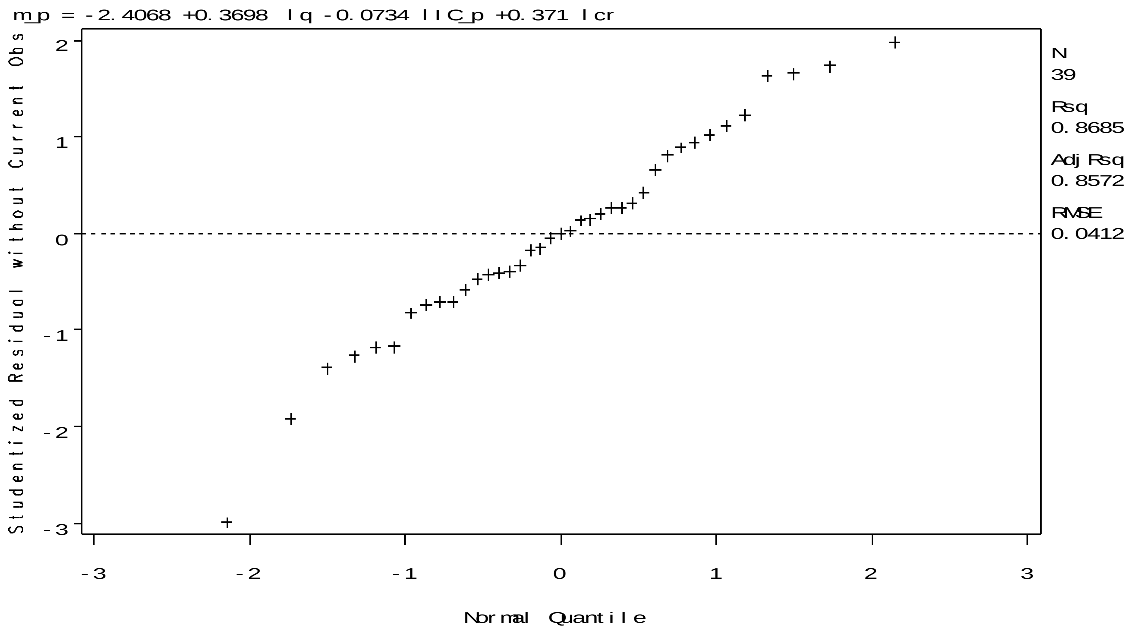


FIGURE A.13 Quantile-Quantile plot for Model (7)

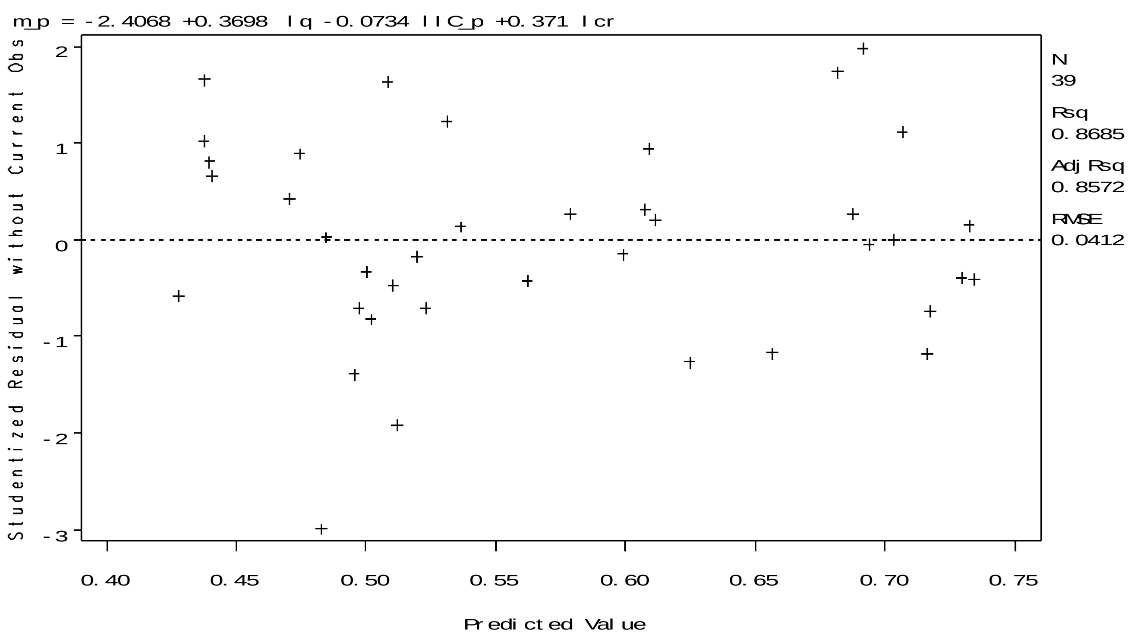


FIGURE A.14 Studentized Residual VS predicted dependent Variable for Model (7)

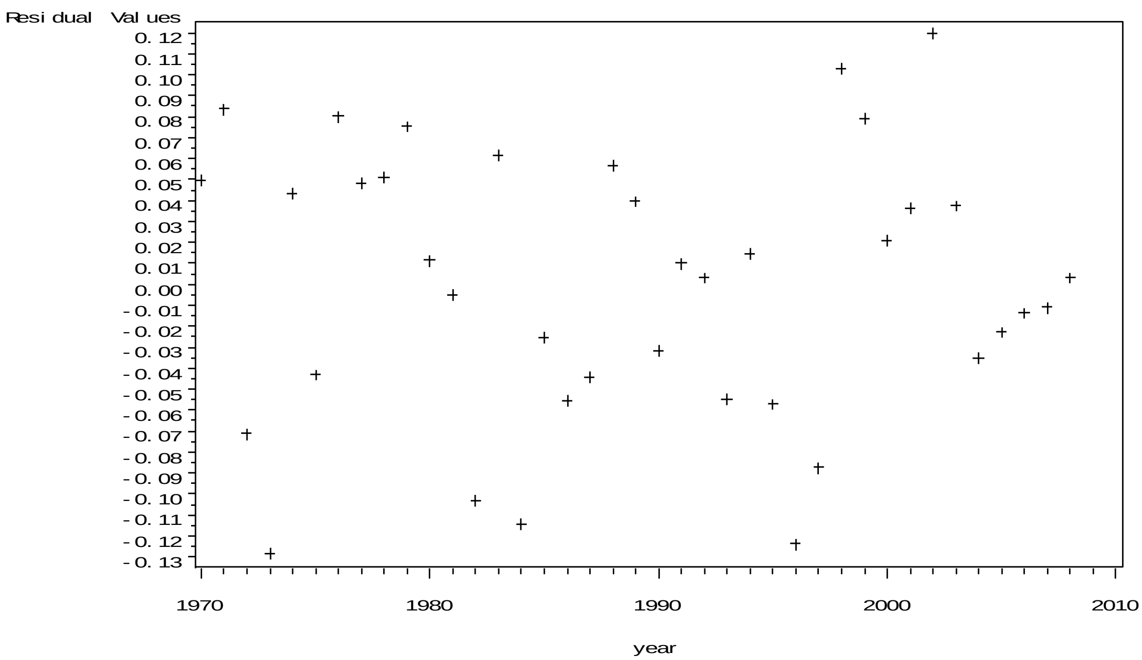


FIGURE A.15 **Studentized Residual VS Year for Model (5)**

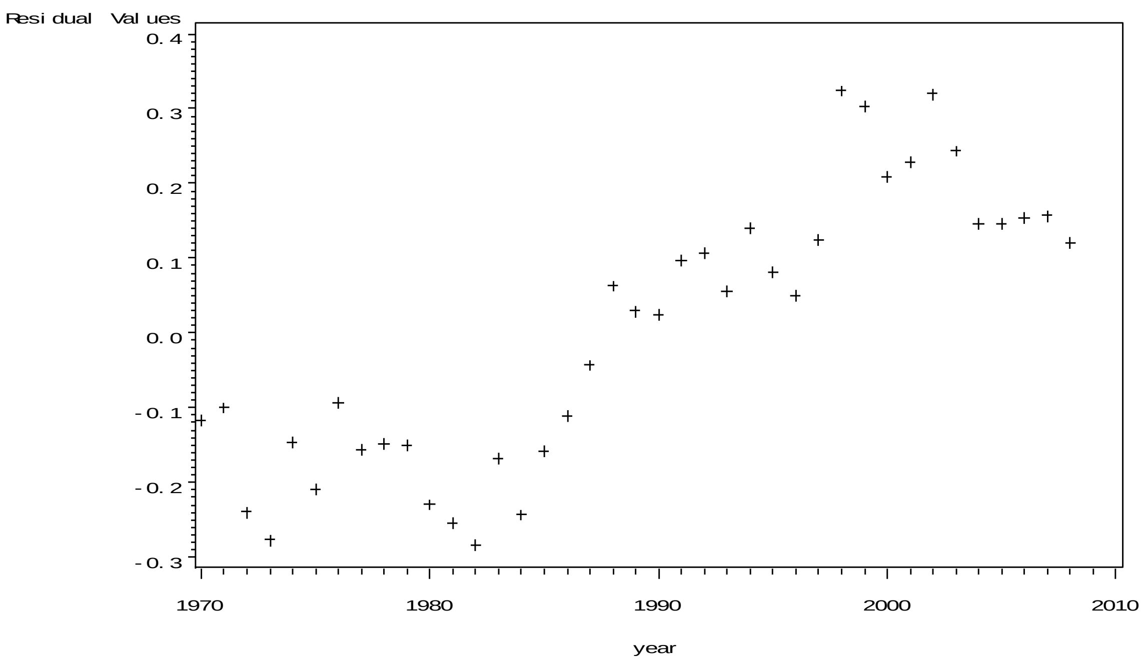
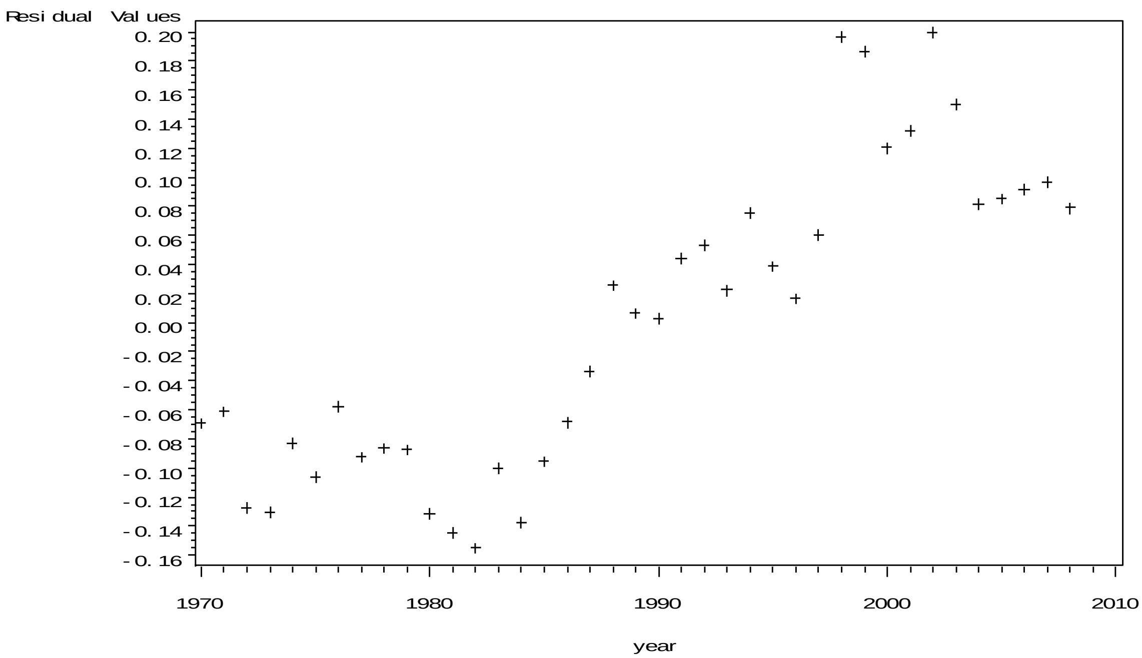


FIGURE A.16 **Studentized Residual VS Year for Model (6)**



<u>FIGURE A.17</u> Studentized Residual VS Year for Model (7)

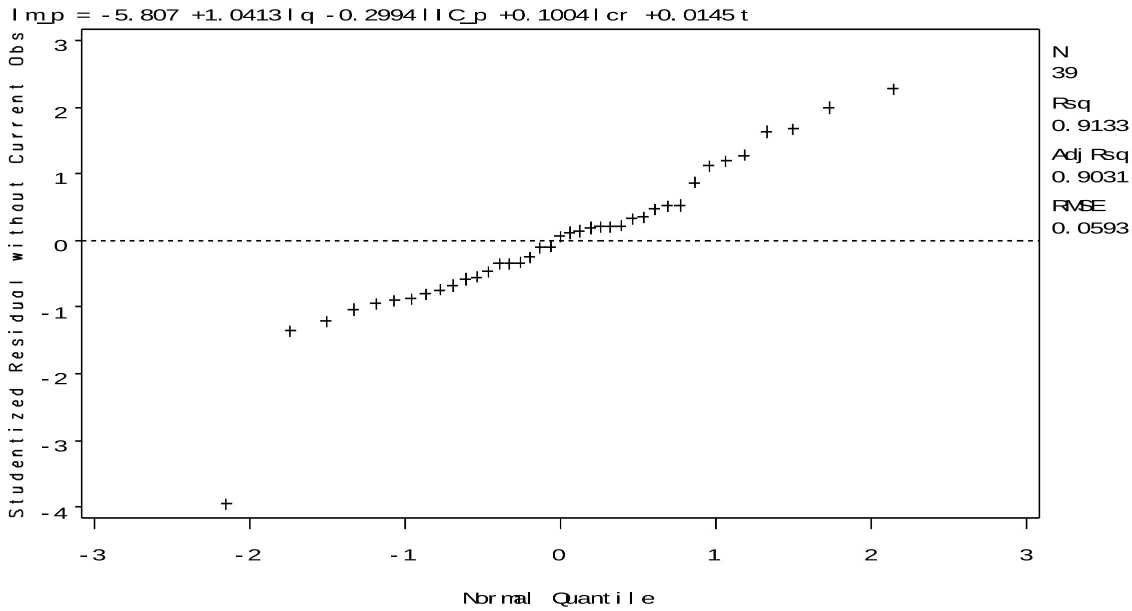


FIGURE A.18 **Quantile-Quantile plot for Model (8)**

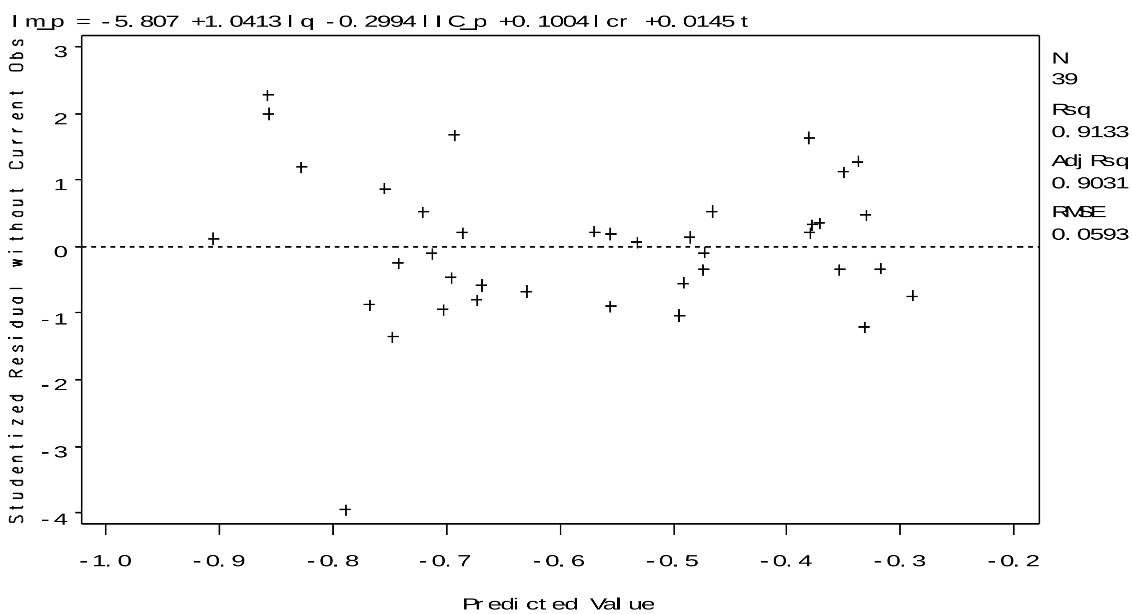


FIGURE A.19 **Studentized Residual VS predicted dependent Variable for Model (8)**

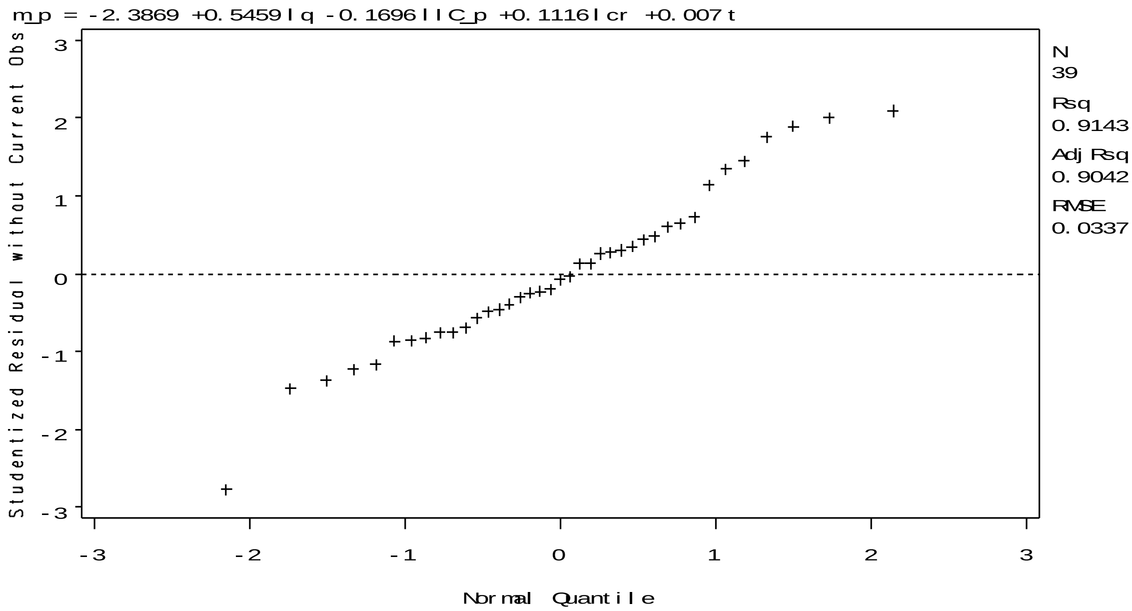


FIGURE A.20 **Quantile-Quantile plot for Model (9)**

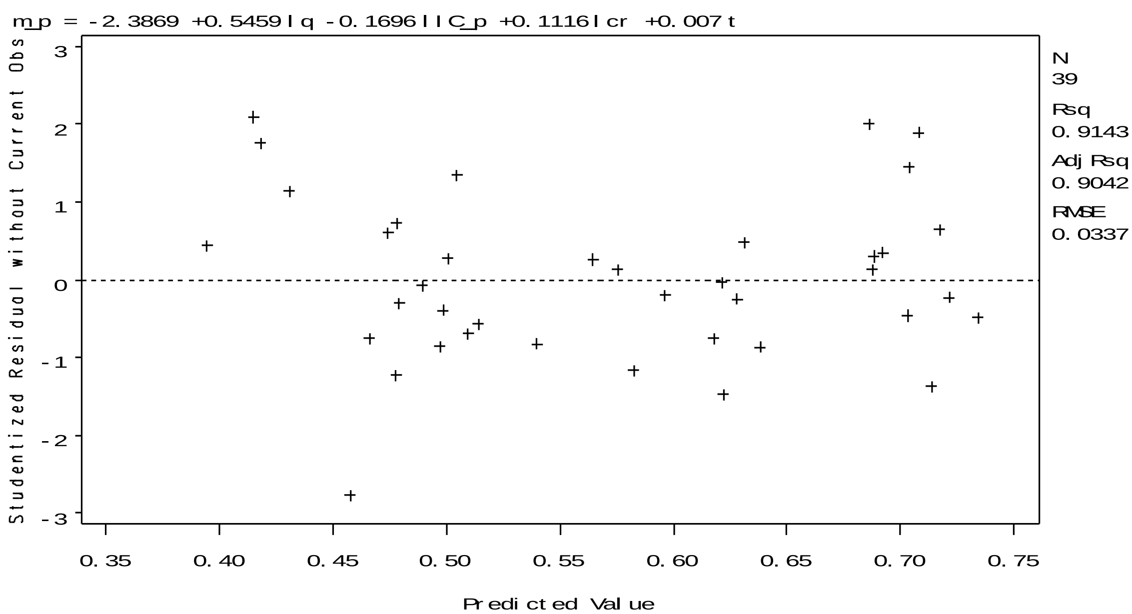


FIGURE A.21 **Studentized Residual VS predicted dependent Variable for Model (9)**

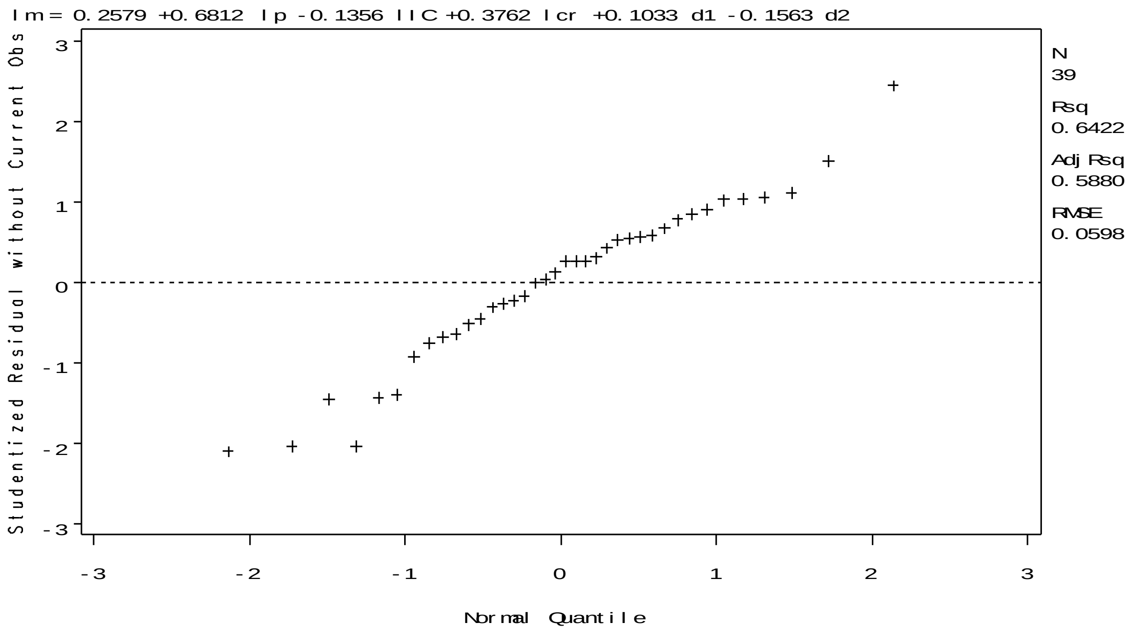


FIGURE A.22 **Quantile-Quantile plot for Model (10)**

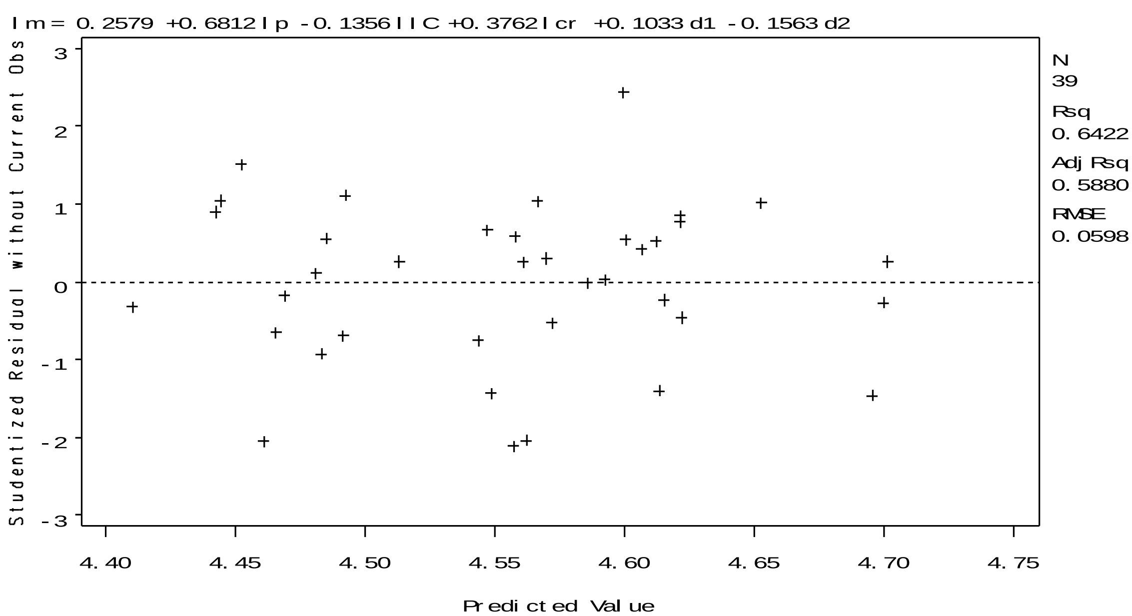
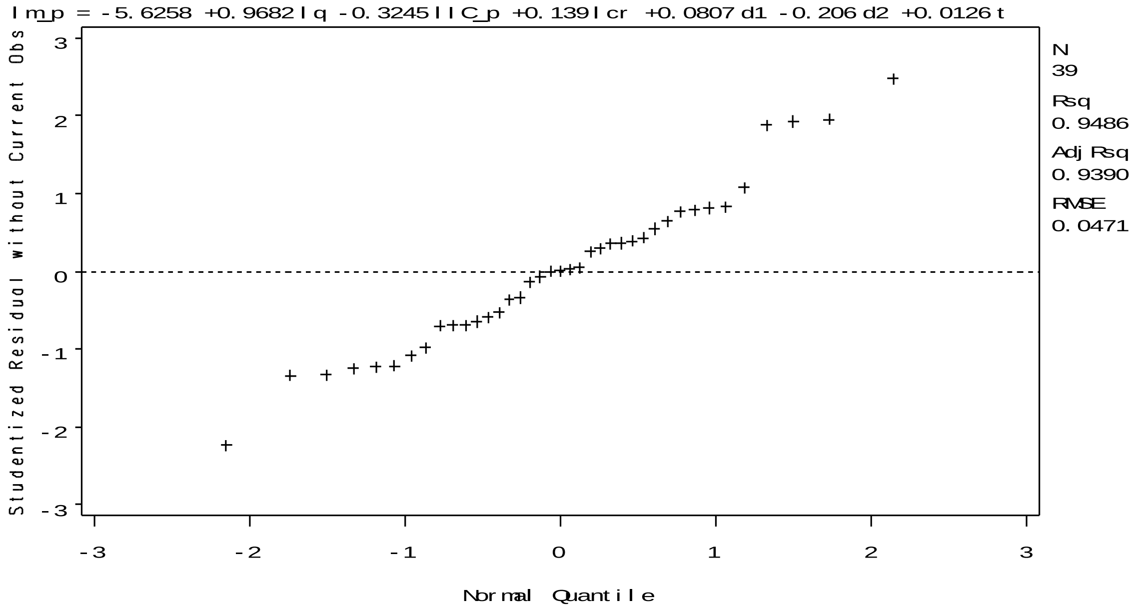
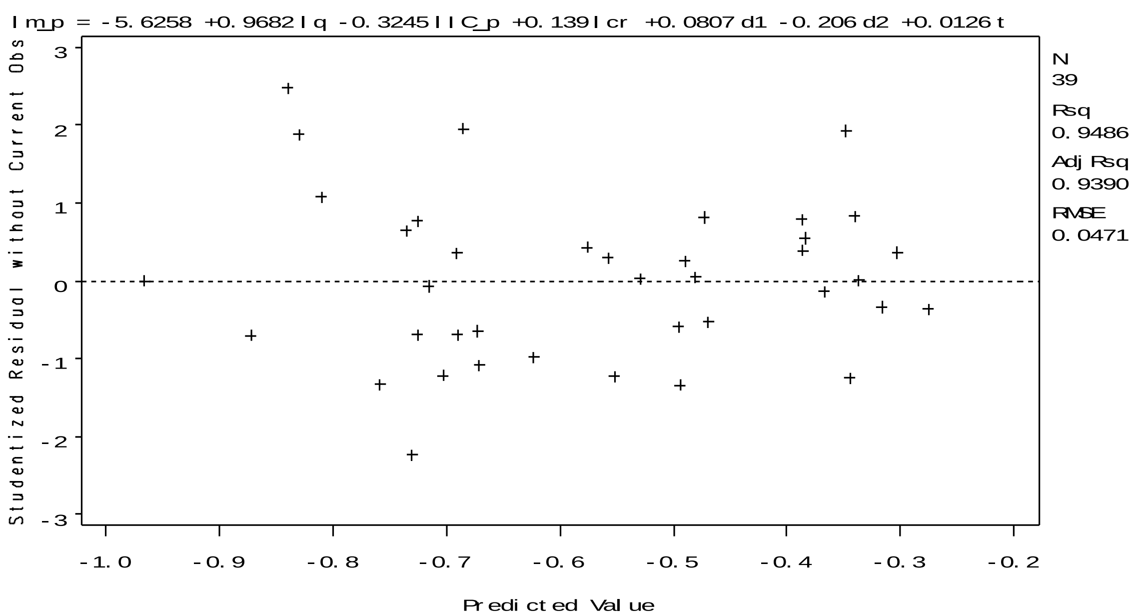


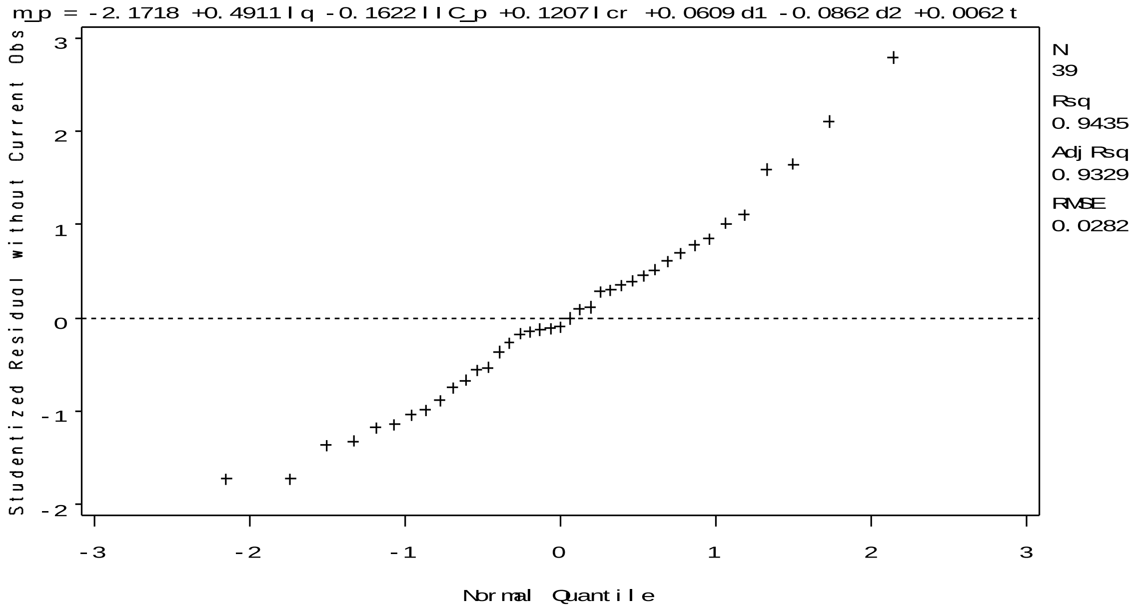
FIGURE A.23 **Studentized Residual VS predicted dependent Variable for Model (10)**



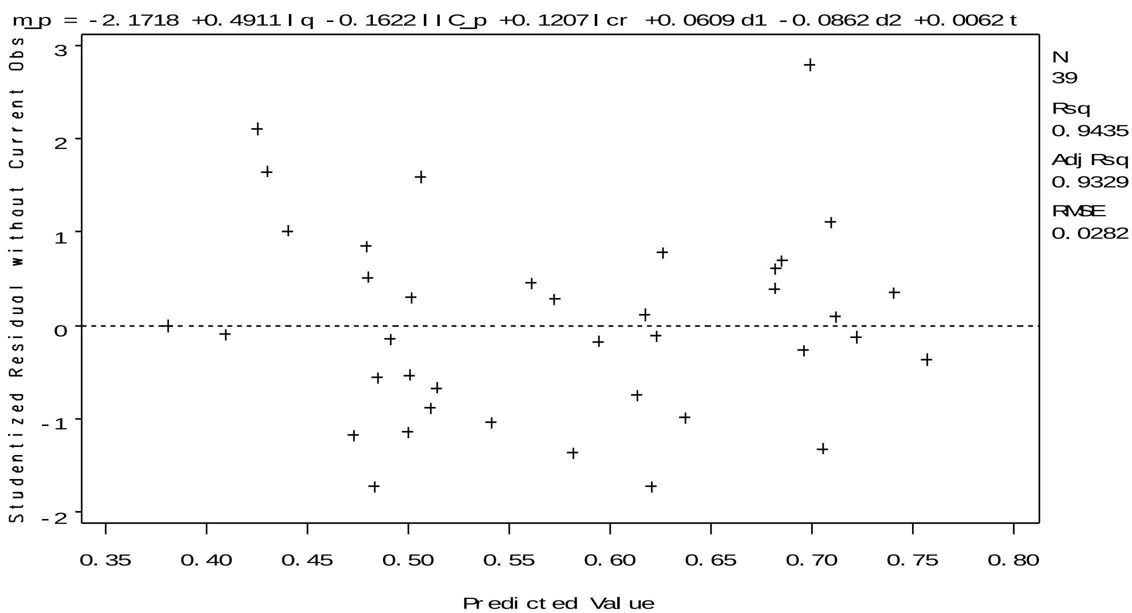
<u>FIGURE A.24</u> **Quantile-Quantile plot for Model (11)**



<u>FIGURE A.25</u> **Studentized Residual VS predicted dependent Variable for Model (11)**



<u>FIGURE A.26</u> **Quantile-Quantile plot for Model (12)**



<u>FIGURE A.27</u> **Studentized Residual VS predicted dependent Variable for Model (12)**