The Interaction Effects of Family Aid Programs on Child Mathematics Achievement –

A Quantitative Study of Women, Infant and Children (WIC) Nutrition Program

Past literatures proved that nutrition has a meaningful impact on the development of children. To support families and children, Women, Infant and Children (WIC) Nutrition Program was started by the government to improve child development. This research examines if the impact of WIC program on children's mathematics achievement is moderated by family income, race, and the current age of the child. The results of this research have important policy implications because they offer an evaluation of the current subsidy program and suggestions to future programs

**Methods**

Participants of this study include 3563 children from 3 to 13 years old. They come from families with total annual income in 1997 from $-72296.16 to $784610.59. Among these children, 1440 of them participated in the WIC program. All children were given a test on Woodcock-Johnson Mathematics Achievement as a measurement of intellectual development, which has a minimum score of 0 and a maximum of 98. The definitions of investigated variables are listed below.

| | |
|---|---|
| **WICpreg** | Women, Infant and Children (WIC) Nutrition Program participant during pregnancy: 0 = No, 1 = Yes. |
| **mathraw97** | Woodcock-Johnson Revised Mathematics Achievement Test Raw Score. Minimum = 0, Maximum = 98. |
| **AGE97** | The child's age in 1997. Minimum = 3, Maximum = 13 |
| **faminc97** | Total family income in 1997 (in 2002 constant dollars). Minimum = $-72296.26, Maximum = $784610.59. |
| **Home97** | A composite total score of the emotional and cognitive stimulation at home. Minimum = 7, Maximum = 27 |
| **RACE** | Centered Binary Coding of Race ( -0.5 = Black, 0.5 = White) |

Independent variables in this study include WICpreg, AGE97, faminc97, Home97, and RACE. The methods of analysis are as follows. This study will first diagnose each assumptions of the general linear model, examine outliers, and transform variables to avoid assumption violation. Then this study will create new centered variables to examine the interaction effects between income and WIC, race and WIC, as well as the current age of the child with WIC. Finally, the results section will present the results of multiple regression analysis before and after the interaction effects are added to the model as well as discuss the model differences.

**Results**

Table 1 covers the numbers, means, medians, proportions, standard deviations, and ranges for each of the investigated variables, which include WIC participation (WICpreg), mathematic test scores (mathraw97), age (AGE97), family income (faminc97), a composite total score of the emotional and cognitive stimulation at home (HOME97), and race (RACE).

**TABLE 1.**
NUMBERS, MEANS, MEDIANS, PROPORTIONS, STANDARD DEVIATIONS, AND RANGES FOR ALL INVESTIGATED VARIABLES

| | WICpreg | mathraw97 | AGE97 | faminc97 | HOME97 | RACE |
|---|---|---|---|---|---|---|
| N | 3322 | 2211 | 2223 | 3563 | 3563 | 3322 |
| mean | 0.43 | 36.33 | 7.47 | 49841.25 | 18.92 | 0.03 |
| median | 0 | 37 | 7 | 39118.44 | 19.2 | 0.5 |
| 0% | 0 | 0 | 3 | -72296.26 | 7 | -0.5 |
| 25% | 0 | 15 | 5 | 20175.7 | 16 | -0.5 |
| 50% | 0 | 37 | 7 | 39118.44 | 19.2 | 0.5 |
| 75% | 1 | 54 | 10 | 64494.99 | 21.8 | 0.5 |
| 100% | 1 | 98 | 13 | 784610.59 | 27 | 0.5 |
| sd | 0.5 | 22.3 | 2.9 | 49751.1 | 3.6 | 0.5 |
| minimum | 0 | 0 | 3 | -72296.26 | 7 | -0.5 |
| maximum | 1 | 98 | 13 | 784610.59 | 27 | 0.5 |

Table 2 provides the frequencies for WIC participation, with 1=yes and 0=no. Among all the children in the study, 1440 of them participated in the WIC program and 1882 of them did not. Table 3 presents correlations among the investigated variables.

**TABLE 2.**
FREQUENCIES FOR WIC PARTICIPATION

| | 0 | 1 |
|---|---|---|
| tableWICpreg | 1882 | 1440 |

**TABLE 3.**
CORRELATION MATRIX FOR ALL INVESTIGATED VARIABLES

| | WICpreg | mathraw97 | AGE97 | faminc97 | HOME97 | RACE |
|---|---|---|---|---|---|---|
| WICpreg | - | | | | | |
| mathraw97 | -0.19 | - | | | | |
| AGE97 | -0.09 | 0.92 | - | | | |
| faminc97 | -0.39 | 0.16 | 0.05 | - | | |
| HOME97 | -0.42 | 0.31 | 0.19 | 0.4 | - | |
| RACE | -0.49 | 0.11 | -0.02 | 0.36 | 0.45 | - |

Assumptions of the general linear model, including linearity, homoscedasticity, and normality of residuals, are diagnosed. For the linearity assumption, AGE97 and faminc97 do not have linear relationships with the dependent variable mathraw97. Therefore, faminc97 was log-transformed into logfaminc, and AGE97 was centered to produce AGE97c, squared AGE97c2, and cubed AGE97c3. The factors of discrepancy, leverage, and influence were examined to determine the outliers of the model. Based on the result of cook's d, 19 observations were omitted from the sample. After the outliers were removed, the model residuals are normally distributed, and the variance of the residuals is constant. Thus, the final model adheres to the linear assumptions of homoscedasticity and normality. The model does not have the problem of multicollinearity.

To examine if the effect of WIC program participation during pregnancy is moderated by family income, race, and the current age of the child, new centered variables were created for all of the continuous variables. HOME97 was centered to produce chome and logfaminc was centered to produce cincome. cincomeWIC and RACEWIC were created to examine the moderator effect of income and race. AGE97cWIC, AGE97c2WIC and AGE97c3WIC were created to analyze the moderator effect of age. All investigated variables in the final model are listed below.

| | |
|---|---|
| **WICpreg** | Women, Infant and Children (WIC) Nutrition Program participant during pregnancy: 0 = No, 1 = Yes. |
| **mathraw97** | Woodcock-Johnson Revised Mathematics Achievement Test Raw Score. Minimum = 0, Maximum = 98. |
| **AGE97c** | The centered child's age in 1997. |
| **AGE97c2** | The squared centered child's age in 1997. |
| **AGE97c3** | The cubed centered child's age in 1997. |
| **cincome** | Centered log total family income in 1997 (in 2002 constant dollars). |
| **chome** | Centered composite total score of the emotional and cognitive stimulation at home. |
| **RACE** | Centered Binary Coding of Race ( -0.5 = Black, 0.5 = White) |
| **cincomeWIC** | To measure the interaction effect of family income and WIC program. |
| **RACEWIC** | To measure the interaction effect of race and WIC program. |
| **AGE97cWIC**<br>**AGE97c2WIC**<br>**AGE97c3WIC** | To measure the interaction effect of age and WIC program. |

The regression result for the main effects model is presented in table 4. There are in total 1829 observations and all independent variables are significant. The intercept tells us that the expected mathematic score is 37.07 for the child who has the average age in the sample, did not participate in the program, and comes from a family with the average income as well as parenting practices in the sample. The parameter estimate for WICpreg is -1.62, meaning that children participating in WIC program on average score about 1.62 points lower than children who did not participate, controlling for age, family income, parenting practices, and race. Regarding other independent variables, the parameter estimate for AGE97 is 8.42, which means that the mathematic score increases on average 8.42 points for each year increase in age, controlling for other independent variables. The effect of the age-squared variable is -0.05 and tells us that this increase in math score decelerates by 0.1 points for each additional year. The effect of AGE97c3 is -0.1, which means that this deceleration decreases by 0.2 points for each additional year.

**TABLE 4.**
MULTIPLE REGRESSION ANALYSIS ON CHILD MATHEMATIC SCORES (mathraw97) (MAIN EFFECTS)

| | **mathraw97** | | |
|---|---|---|---|
| *Predictors* | *Estimates* | *sd* | *p* |
| (Intercept) | 37.07 | 0.32 | **<0.001** |
| WICpreg | -1.62 | 0.43 | **<0.001** |
| AGE97c | 8.42 | 0.15 | **<0.001** |
| AGE97c2 | -0.05 | 0.02 | **0.042** |
| AGE97c3 | -0.10 | 0.01 | **<0.001** |
| cincome | 0.49 | 0.17 | **0.005** |
| RACE | 2.37 | 0.43 | **<0.001** |
| chome | 0.58 | 0.07 | **<0.001** |
| Observations | 1829 | | |
| $R^2$ / $R^2$ adjusted | 0.8889 / 0.8885 | | |

For cincome, the parameter estimate is 0.49, which means that an increase in family income by 1 percent increases mathematic scores by 0.0049, holding other independent variables constant. For race, the parameter estimate is 2.37, indicating that white children on average score 2.37 points higher than black children, holding other independent variables constant. For chome, the parameter estimate is 0.58, suggesting that a child will score on average 0.58 points higher for every additional point in parenting practices, holding all other independent variables constant. The total variation of child mathematic achievement accounted for by the main effects model is 88.89%, and 88.85% after being adjusted.

The result of the multiple regression analyses with interaction effects of family income and WIC program is presented in table 5. It shows that the p-value for the cincome*WICpreg interaction term is significant, meaning that the impact of WIC program on children's mathematic scores depends on the level of family income. The result indicates that on average, children in WIC program have mathematic scores that are 1.76 points lower than children who are not in WIC program. However, this disparity is not uniform across income levels. For every 1-unit increase in the logged family income, a child not participating in the WIC program should expect to earn an additional 1.04 points, while a child participating in the WIC program should only expect to earn an additional 0 point. The total variation of child mathematic achievement accounted for by the interaction model is 88.96%, and 88.91% after being adjusted. The fact that $R^2$ / $R^2$ adjusted in this model is higher than the $R^2$ / $R^2$ in the main effects model shows that taking interaction effects into account increase the explanatory power of this model.

The result of the multiple regression analysis with interaction effects of race and WIC program is presented in table 6, which shows that the p-value for the RACE*WICpreg interaction term is significant, meaning that there is evidence for WICpreg being moderated by RACE. In the model, the parameter estimate for WICperg is -1.88, suggesting that on average, children participating in the WIC program have mathematic scores 1.88 points lower than children who did not participate in the WIC program. However, such disparity is not constant for different racial groups. White children in the WIC program should expect to earn an additional 0.345 points (3.41*0.5 – 2.72*0.5), while black children in WIC should expect to lose 0.345 points (3.41*(-0.5) – 2.72*(-0.5)). The total variation of child mathematic achievement accounted for by the interaction model is 88.96%, and 88.91% after being adjusted. More total variation was accounted for by this model because taking the moderator effects of race into account increases the explanatory power of the model.

The result of the multiple regression analysis with interaction effects of age and WIC program is presented in table 7. While the p-values for the AGE97c*WICpreg and AGE97c3*WICpreg interaction terms are significant, the p-value for AGE97c2*WICpreg is not significant. The parameter estimate for WICpreg is -1.57, which means that on average, children participating in the WIC program scores 1.57 lower in mathematic tests than children who did not participate in the program. However, this disparity is not the same across age. For every 1-unit increase in children's age, children participating in the WIC program should expect to earn an additional 7.92 points. The total variation of child mathematic achievement accounted for by the interaction model is 88.99%, and 88.92% after being adjusted, meaning that taking the moderator effect of age into account increases the explanatory power of our model.

**TABLE 5.**

MULTIPLE REGRESSION ANALYSIS ON CHILD MATHEMATIC SCORES (mathraw97)
(INTERACTION EFFECTS ON FAMILY INCOME AND WIC PROGRAM)

| | **mathraw97** | | |
|---|---|---|---|
| *Predictors* | *Estimates* | *sd* | *p* |
| (Intercept) | 36.86 | 0.32 | **<0.001** |
| WICpreg | -1.76 | 0.44 | **<0.001** |
| AGE97c | 8.43 | 0.15 | **<0.001** |
| AGE97c2 | -0.05 | 0.02 | **0.050** |
| AGE97c3 | -0.10 | 0.01 | **<0.001** |
| cincome | 1.04 | 0.24 | **<0.001** |
| RACE | 2.26 | 0.43 | **<0.001** |
| chome | 0.58 | 0.07 | **<0.001** |
| cincomeWIC | -1.04 | 0.33 | **0.001** |
| Observations | 1829 | | |
| $R^2$ / $R^2$ adjusted | 0.8896 / 0.8891 | | |

**TABLE 6.**

MULTIPLE REGRESSION ANALYSIS ON CHILD MATHEMATIC SCORES (mathraw97)
(INTERACTION EFFECTS ON RACE AND WIC PROGRAM)

| | **mathraw97** | | |
|---|---|---|---|
| *Predictors* | *Estimates* | *sd* | *p* |
| (Intercept) | 36.83 | 0.33 | **<0.001** |
| WICpreg | -1.88 | 0.44 | **<0.001** |
| AGE97c | 8.44 | 0.15 | **<0.001** |
| AGE97c2 | -0.05 | 0.02 | **0.053** |
| AGE97c3 | -0.10 | 0.01 | **<0.001** |
| cincome | 0.48 | 0.17 | **0.005** |
| RACE | 3.41 | 0.53 | **<0.001** |
| chome | 0.57 | 0.07 | **<0.001** |
| RACEWIC | -2.72 | 0.83 | **0.001** |
| Observations | 1829 | | |
| $R^2$ / $R^2$ adjusted | 0.8896 / 0.8891 | | |

**TABLE 7.**
MULTIPLE REGRESSION ANALYSIS ON CHILD MATHEMATIC SCORES (mathraw97)
(INTERACTION EFFECTS ON AGE AND WIC PROGRAM)

| | **mathraw97** | | |
|---|---|---|---|
| *Predictors* | *Estimates* | *sd* | *p* |
| (Intercept) | 37.03 | 0.36 | **<0.001** |
| WICpreg | -1.57 | 0.60 | **0.009** |
| AGE97c | 8.75 | 0.19 | **<0.001** |
| AGE97c2 | -0.05 | 0.03 | 0.131 |
| AGE97c3 | -0.11 | 0.01 | **<0.001** |
| cincome | 0.51 | 0.17 | **0.003** |
| RACE | 2.36 | 0.43 | **<0.001** |
| chome | 0.59 | 0.07 | **<0.001** |
| AGE97cWIC | -0.83 | 0.31 | **0.007** |
| AGE97c2WIC | -0.01 | 0.05 | 0.804 |
| AGE97c3WIC | 0.03 | 0.19 | 0.181 |
| Observations | 1829 | | |
| $R^2$ / $R^2$ adjusted | 0.8899 / 0.8892 | | |

**Conclusion**

This study examines the moderator effect of family income, race, and age on WIC program through multiple regression analysis, and the results show that all three of them had a significant impact on WIC program. Family income has a negative interaction effect with WIC program and the reason may be that children's scores are negatively impacted by WIC. In terms of race, White children participating in the WIC program earn more points than black children, which is another reflection of the impact of racial inequality on children's academic performance. Older children participating in the WIC program earn more scores than younger children given that children's intellectual capability increases with age. Policymakers are recommended to take the moderators into account when designing future programs. The subsidies of the program should be adjusted to target children of diverse racial, ethnic, and age group with different financial background. Only in this way can the program meet the diverse needs from children.

Appendix

```
good<-read.csv("/Users/yunxin/Desktop/PENN/Courses/Linear Modeling/Assignment 
3/good.csv")
good<-good[ ,c("WICpreg","mathraw97","AGE97","faminc97","HOME97","CHRACE")]
good$RACE <- ifelse(good$CHRACE== 9, NA, ifelse(good$CHRACE== 1, .5, 
ifelse(good$CHRACE== 2, -.5, NA)))
goodR<-good[ ,c("WICpreg","mathraw97","AGE97","faminc97","HOME97","RACE")]

# descriptive statistics
N<-
c(length(which(!is.na(goodR$WICpreg))),length(which(!is.na(goodR$mathraw97))),length(whic
h(!is.na(goodR$AGE97))),length(which(!is.na(goodR$faminc97))),length(which(!is.na(goodR$
HOME97))),length(which(!is.na(goodR$RACE))))
mean<-sapply(goodR,mean,na.rm=TRUE)
median<-sapply(goodR,median,na.rm=TRUE)
quantile<-sapply(goodR,quantile,na.rm=TRUE)
sd<-sapply(goodR,sd,na.rm=TRUE)
range<-sapply(goodR,range,na.rm=TRUE)
descriptive<-rbind(N,mean,median,quantile,sd,range)
write.csv(descriptive,"/Users/yunxin/Desktop/PENN/Courses/Linear Modeling/Assignment 
3/descriptive.csv")
tableWICpreg<-table(goodR$WICpreg)
frequency<-tableWICpreg
write.csv(frequency,"/Users/yunxin/Desktop/PENN/Courses/Linear Modeling/Assignment 
3/frequency.csv")
cor<-cor(goodR,use = "complete.obs")
correlation<-round(cor,2)
write.csv(correlation,"/Users/yunxin/Desktop/PENN/Courses/Linear Modeling/Assignment 
3/correlation.csv")

goodR2<-
na.omit(goodR[ ,c("WICpreg","mathraw97","AGE97","faminc97","RACE","HOME97")])
# assumptions
# linearity
pairs(goodR2,panel=panel.smooth)
library(ggplot2)

ggplot(goodR2,aes(x=AGE97, y = mathraw97)) +
  geom_point(size = 0.6) +
  xlab("AGE97") +
  ylab("mathraw97") +
  theme_bw() +
  geom_smooth(method = "loess")
ggplot(goodR2,aes(x=faminc97, y = mathraw97)) +
  geom_point(size = 0.6) +
```

```
  xlab("faminc97") +
  ylab("mathraw97") +
  theme_bw() +
  geom_smooth(method = "loess")
ggplot(goodR2,aes(x=HOME97, y = mathraw97)) +
  geom_point(size = 0.6) +
  xlab("HOME97") +
  ylab("mathraw97") +
  theme_bw() +
  geom_smooth(method = "loess")

#homoscedasticity
lm<-lm(mathraw97~WICpreg+AGE97+faminc97+HOME97+RACE,data=goodR2)
goodR2.res<-resid(lm)
fitted.res<-fitted(lm)
plot(fitted.res,goodR2.res)
abline(0,0,col="red")
lines(lowess(goodR2.res~fitted.res),col="green")

#normality of residuals
hist(goodR2.res,15)

#outliers
lm2<-lm(mathraw97~WICpreg+AGE97+faminc97+HOME97+RACE,data=goodR2)
outliers<-goodR2[,c("mathraw97","WICpreg","AGE97","faminc97","RACE","HOME97")]
outliers$r<-rstudent(lm2)
outliers$lev<-hatvalues(lm2)
outliers$cd<-cooks.distance(lm2)
outliers$dffit<-dffits(lm2)
dfbetaR<-dfbetas(lm2)
colnames(dfbetaR)
colnames(dfbetaR)<-
c("int_dfb","WICpreg_dfb","AGE97_dfb","faminc97_dfb","HOME97_dfb","RACE_dfb")
outliers<-cbind(outliers,dfbetaR)
head(outliers)
# discrepancy
plot(outliers$r,
     xlab="Index",
     ylab="studentized residuals",
     pch=19)
abline(h = 0, col = "red")
abline(h = -3, col = "green")
abline(h = 3, col = "green")
abline(h = -2, col = "blue")
abline(h = 2, col = "blue")
rstudent1<- subset(outliers,abs(r)>3)
```

```
# leverage
# (2*5+2)/2042=0.0059
plot(outliers$lev,
    xlab="Index",
    ylab="leverage",
    pch=19)
abline(h=0.0059, col="blue")
leverage1<- subset(outliers,lev >0.01)
View(leverage1)
# influence - Cook's D
plot(outliers$cd,
    xlab="Index",
    ylab="cook’s d",
    pch=19)
abline(h=4/1848, col="blue")
large_cd<-subset(outliers, cd > (4/1848))
quantile(large_cd$cd, probs = seq(0, 1, 0.025))
abline(h= .0057, col="green")
to_omit <- subset(outliers, cd >= .0057)
to_omit
#multicolinearity
library(car)
multicolinearity<-vif(lm2)
multicolinearity
#no multicolinearity

#remove outliers
no_outliers<-subset(outliers, cd < .0057)

#transform faminc97
min(no_outliers$faminc97)
no_outliers$logfaminc<-ifelse(no_outliers$faminc97<=
1,0,ifelse(no_outliers$faminc97>1,log(no_outliers$faminc97), NA))
#transform AGE97
no_outliers$AGE97c<-no_outliers$AGE97-mean(no_outliers$AGE97)
no_outliers$AGE97c2<-(no_outliers$AGE97c)^2
no_outliers$AGE97c3<-(no_outliers$AGE97c)^3

#examine homoscedasticity again
lm<-
lm(mathraw97~WICpreg+AGE97c+AGE97c2+AGE97c3+logfaminc+HOME97+RACE,data=n
o_outliers)
no_outliers.res<-resid(lm)
fitted.res<-fitted(lm)
plot(fitted.res,no_outliers.res)
abline(0,0,col="red")
```

```
lines(lowess(no_outliers.res~fitted.res),col="green")

#examine normality of residuals again
hist(no_outliers.res,15)

#data transformation

#new centered variables
no_outliers$chome<-no_outliers$HOME97 - mean(no_outliers$HOME97)
no_outliers$cincome<-no_outliers$logfaminc - mean(no_outliers$logfaminc)
#interaction variables
no_outliers$cincomeWIC<-no_outliers$cincome * no_outliers$WICpreg
no_outliers$RACEWIC<-no_outliers$RACE * no_outliers$WICpreg
no_outliers$AGE97cWIC<-no_outliers$AGE97c * no_outliers$WICpreg
no_outliers$AGE97c2WIC<-no_outliers$AGE97c2 * no_outliers$WICpreg
no_outliers$AGE97c3WIC<-no_outliers$AGE97c3 * no_outliers$WICpreg

library(lmSupport)
library(sjPlot)
#main effects model
lm_main<-lm(mathraw97 ~ WICpreg + AGE97c + AGE97c2 + AGE97c3+ cincome + RACE +
chome, data=no_outliers)
summary(lm_main)
lm_main<-tab_model(lm_main)
lm_main

#interaction between WIC and income
lm_inter_1<-lm(mathraw97 ~ WICpreg + AGE97c + AGE97c2 + AGE97c3 + cincome + RACE
+ chome + cincomeWIC, data=no_outliers)
summary(lm_inter_1)
lm_inter_1<-tab_model(lm_inter_1)
lm_inter_1

#interaction between WIC and race
lm_inter_2<-lm(mathraw97 ~ WICpreg + AGE97c + AGE97c2 + +AGE97c3 + cincome +
RACE + chome + RACEWIC, data=no_outliers)
summary(lm_inter_2)
lm_inter_2<-tab_model(lm_inter_2)
lm_inter_2

#interaction between WIC and age
lm_inter_3<-lm(mathraw97 ~ WICpreg + AGE97c + AGE97c2 + AGE97c3 + cincome + RACE
+ chome + AGE97cWIC +AGE97c2WIC + AGE97c3WIC, data=no_outliers)
summary(lm_inter_3)
lm_inter_3<-tab_model(lm_inter_3)
lm_inter_3
```