Term Paper

The Impact of Smoking and Alcohol on the Birth Weight

Table of Contents

Abstract

The main paper of this paper is wanted to examine the effect of mothers' smoking and alcohol during pregnancy on birth weight. For this purpose, take the data with the name "BWGHT2" which consists of 23 variables and the sample size of the data is 1832 observations. Two regression models are used in this study which concludes that there has a negative effect of smoking cigarettes and drinking alcohol of a mother during pregnancy on birth weight. The models also show that there has a positive effect of the father's age, five-minute apgar score, male baby, black and white father, and negative effect of black mother on birth weight.

Introduction

This paper used data with the name "BWGHT2" which consists of 23 variables and the sample size of the data is 1832 observations. This study used 9 variables consisting of one independent, two explanatory, and six relevant control variables. The previous paper's description is given below.

David H. Robin et al (1986) examine the effect of passive smoking on birth weight. They took the data of 500 consecutive Danish women and then applied a multiple linear regression model. This study concluded that birth weight was found to be reduced by maternal smoking exposure, and the father's indirect or passive smoking exposure had an effect that was almost as significant (66%) as the mothers. On average, each pack of cigarettes smoked by the father each day resulted in a 120 g reduction in birth weight. After adjusting for the mother's age, parity, alcohol and cigarette use during pregnancy, illness during pregnancy, social status, and the sex of the baby, this relationship still had statistical significance. The lower social classes felt the effects of passive smoking the most.

The Data

The data is selected with the name "BWGHT2" which consists of 23 variables and the sample size of the data is 1832 observations. This study used 9 variables consisting of one independent, two explanatory, and six relevant control variables, which include birth weight, average cigarettes per day, average drinks per week, father's age, five-minute apgar score, male baby, black mother, black father, and white father. The Gretl output of descriptive statistics is given below.

Table 1: Summary Statistics

Variable	Mean	Median	S.D.	Min	Max
fage	31.9	31	5.71	18	64
bwght	3.40E+03	3.43E+03	577	360	5.20E+03
fmaps	9	9	0.48	2	10
cigs	1.09	0	4.22	0	40
drink	0.0198	0	0.289	0	8
male	0.514	1	0.5	0	1
mblck	0.0595	0	0.237	0	1
fwhte	0.89	1	0.313	0	1
fblck	0.0584	0	0.235	0	1

Econometric Model

For the purpose of checking the impact of a mother's cigarette smoking and drinking alcohol during pregnancy period on the birth weight using a regression model. There is one dependent variable "birth weight" and more than one independent or explanatory variables "cigarette smoking and drinking alcohol" so we use a multiple regression model but the dependent variable is in quantitative form as well as the independent variable in the original form which means that there is a linear relationship between dependent and independent variables so we apply multiple linear regression model. There is a lot of estimation technique of the regression model such as ordinary least square (OLS), maximum likelihood (MLE), and method of moment (MOM), but for the linear regression model ordinary least square estimation technique is one of best estimation techniques because the aim of OLS is to minimize the residual. Use two multiple linear regression models in this study, in first model use birth weight as a dependent, and average cigarettes per day, and average drinks per week are used as explanatory variables. This model formal equation can be written as

$$BWGHT_i = \beta_0 + \beta_1 CIGS_i + \beta_2 DRINK_i + ui \qquad (1)$$

Here Bwght is the dependent variable "Birth Weight", and Cigs and Drink are explanatory variables "average cigarettes per day, and average drinks per week" respectively. β_0 is intercept of the model, β_1, and β_2 are the slope parameter of the model. Use birth weight as a dependent, and average cigarettes per day, and average drinks per week are used as explanatory variables as well as use six significant control variables in the second model. The formal equation of the second model can be written as.

$$BWGHT_i = \beta_0 + \beta_1 CIGS_i + \beta_2 DRINK_i + \beta_3 FAGE_i + \beta_4 FMAPES_i + \beta_5 MALE_i + \beta_6 MBLCK_i + \beta_7 FWHTE_i + \beta_8 MBLCK_i + u_i \quad (2)$$

Here Fage, Fmapes, Male, Mblck, Fwhte, and Mblck are used as control variables that have a significant effect on birth weight.

Results

The Gret output of model (1) in the table form is given below.

Table 2: Regression Model (1)

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	3422.79	14.2035	241	<0.0001	***
cigs	−11.9718	3.39967	−3.521	0.0004	***
drink	−8.60616	48.362	−0.1780	0.8588	

Mean dependent var	3409.934	S.D. dependent var	571.159
Sum squared resid	5.55E+08	S.E. of regression	569.318
R-squared	0.007595	Adjusted R-squared	0.00644
F(2, 1711)	6.546973	P-value(F)	0.00147
Log-likelihood	−13304.93	Akaike criterion	26615.9
Schwarz criterion	26632.2	Hannan-Quinn	26621.9

The coefficient of Cigs shows that, As one unit increases in average cigarettes per day, birth weight decreases by 11.9718 units, which means that there has a negative effect of smoking cigarettes during pregnancy on birth weight. The P-value of average cigarettes per day is less than the critical value $\alpha = 0.05$, which means that there has a statistically significant effect of smoking cigarettes during pregnancy on birth weight.

The coefficient of drink shows that, As one unit increases in average drinks per week, birth weight decreases by 8.606 units, which means that there has a negative effect of drinking alcohol during pregnancy on birth weight. The P-value of the average drink per week is greater than the critical value α = 0.05, which means that there has a statistically non-significant effect of drinking alcohol during pregnancy on birth weight.

The coefficient of determination (R^2) value is 0.008, which means that the variation in birth weight is explained 0.8% by the variation in smoking cigarettes and drinking alcohol, which means that the overall model is not well-fitted for future prediction. The P-value of the F-test is less than the critical value α = 0.05, which means that the overall model is statistically significant.

The Gret output of model (1) in the table form is given below.

Table 3: Regression Model (2)

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	1036.32	268.93	3.853	0.0001	***
cigs	−11.8670	3.29073	−3.606	0.0003	***
drink	−14.9484	46.4735	−0.3217	0.7478	
fage	6.74739	2.33621	2.888	0.0039	***
fmaps	210.748	27.9459	7.541	<0.0001	***
male	95.0948	26.5429	3.583	0.0003	***
mblck	−227.722	129.559	−1.758	0.049	***
fwhte	240.494	60.141	3.999	<0.0001	***
fblck	428.237	139.335	3.073	0.0021	***

Mean dependent var	3412.826	S.D. dependent var	562.131
Sum squared resid	5.06E+08	S.E. of regression	546.108
R-squared	0.060623	Adjusted R-squared	0.0562
F(8, 1697)	13.68967	P-value(F)	2.17E-19
Log-likelihood	−13168.80	Akaike criterion	26355.6
Schwarz criterion	26404.58	Hannan-Quinn	26373.7

The coefficient of Cigs shows that, As one unit increases in average cigarettes per day, birth weight decreases by 11.867 units, which means that there has a negative effect of smoking cigarettes during pregnancy on birth weight. The P-value of average cigarettes per day is less than the critical value α = 0.05, which means that there has a statistically significant effect of smoking cigarettes during pregnancy on birth weight. The coefficient of drink shows that, As one unit increases in average drinks per week, birth weight decreases by 14.9484 units, which means that there has a negative effect of drinking alcohol during pregnancy on birth weight. The P-value of the average drink per week is greater than the critical value α = 0.05, which means that there has a statistically non-significant effect of drinking alcohol during pregnancy on birth weight.

The other six control variables also interpreted as.

As one unit increases in the father's age, birth weight also increases by 6.747 units, which means that there has a positive effect of the father's age on birth weight. The P-value of the father's age is less than the critical value α = 0.05, which means that there has a statistically significant effect of the father's age on birth weight. When one unit increases in the five-minute apgar score, birth weight also increases by 210.75 units, which means that there has a positive effect of the five-minute apgar score on birth weight. The P-value of the five-minute apgar score is less than the critical value α = 0.05, which means that there has a statistically significant effect of the five-minute apgar score on birth weight. As one unit increases in the male baby, birth weight also increases by 95.095 units, which means that there has a positive effect of the male baby on birth weight. The P-value of the male baby is less than the critical value α = 0.05, which means that there has a statistically significant effect of the male baby on birth weight.

When one unit increases in the black mother, birth weight decreases by 227.72 units, which means that there has a negative effect of black mother on birth weight. The P-value of the black mother is less than the critical value α = 0.05, which means that there has a statistically significant effect of the black mother on birth weight. As one unit increases in the white father, birth weight also increases by 240.49 units, which means that there has a positive effect of the white father on birth weight. The P-value of the white father is less than the critical value α = 0.05, which means that there has a statistically significant effect of the white father on birth weight. As one unit increases in the black father, birth weight also increases by 428.24 units, which means that there has a positive effect of the black father on birth weight. The P-value of the black father is less than the critical value α = 0.05, which means that there has a statistically significant effect of the black father on birth weight.

The coefficient of determination (R^2) value is 0.06, which means that the variation in birth weight is explained 6% by the variation in smoking cigarettes, drinking alcohol, and all relevant control variables, which means that the overall model is not well-fitted for future prediction. The P-value of the F-test is less than the critical value $\alpha = 0.05$, which means that the overall model is statistically significant.

Diagnostic Testing

t-test

To test the effect of cigs and drink on birth weight individually using t-test statistic.

To test the effect of cigs on birth weight by using t-test is given below.

H_0: $\beta_1 = 0$

H_1: $\beta_1 \neq 0$

Level of significance

$\alpha = 0.05$

Test Statistics:

$$t = \frac{\hat{\beta}_1 - \beta_1}{S.E(\hat{\beta}_1)}$$

Computation:

$$t = \frac{-11.867 - 0}{3.291}$$

t = 3.606

Critical Region:

If t calculated value ≥ t table value then we need to reject the null hypothesis

t calculated value = 3.606

t tabulated value = $t_{0.025(1831)} = 1.96$

Conclusion:

3.606 > 1.96 so we reject the null hypothesis which means that there has a statistically significant effect of smoking cigarettes during pregnancy on birth weight.

Now to test the effect of drinking alcohol during pregnancy on birth weight by using t-test is given below.

H_0: $\beta_2 = 0$

H_1: $\beta_2 \neq 0$

Level of significance

$\alpha = 0.05$

Test Statistics:

$$t = \frac{\hat{\beta}_2 - \beta_2}{S.E(\hat{\beta}_2)}$$

Computation:

$$t = \frac{-14.9484 - 0}{46.4735}$$

t = 0.322

Critical Region:

If t calculated value ≥ t table value then we need to reject the null hypothesis

t calculated value = 0.322

t tabulated value = $t_{0.025(1831)} = 1.96$

Conclusion:

0.322 < 1.96 so we cannot reject the null hypothesis which means that there is no effect of drinking alcohol during pregnancy on birth weight.

F-test

To test the significance of overall model by using the following F-test statistic.

$H_{0:}\ \beta_k = 0$

$H_{1:}\ \beta_2 \neq 0$

Level of significance

$\alpha = 0.05$

Test Statistic:

$$F = \frac{(R^2)/k}{(1-R^2)/(n-k-1)}$$

Computation:

$$F = \frac{0.0606/8}{(1-0.0606)/1823}$$

$$F = \frac{0.00758}{0.00052}$$

F = 14.58

Critical Region:

If F calculated value ≥ F table value then we need to reject the null hypothesis.

Conclusion:

F calculated value = 14.58 and F table value = $F_{0.05(8,1823)}$ = 1.95

14.58 > 1.95 so we reject the null hypothesis which means that the overall model is statistically significant at 5% level of significance.

Heteroscedasticity

Test the assumption of homoscedasticity in the model using Breusch-Pagan test, which is given below.

Null hypothesis: No heteroscedasticity

Chi2 (1)	P-value
55.57	0.000

The P-value of BP test is less than the critical value 0.05, so we reject the null hypothesis, which means that there is the problem of heteroscedasticity in the model.

Specification Bias

To test the omitted variable bias in the model by using Ramsey RESEST test which is given below.

Ho: model has no omitted variables

F (3,1694)	P-value
6.76	0.0002

The P-value of the test is less than the critical value, so we reject the null hypothesis, which means that the model has some omitted variables which are not included in the model.

Conclusion

The first and second regression models conclude that there has a negative effect of smoking cigarettes during pregnancy on birth weight and there is also a negative effect of drinking alcohol during pregnancy on birth weight but there is a non-significant effect of drinking alcohol on birth weight. The models also show that there has a positive effect of the father's age, five-minute apgar score, male baby, black and white father, and negative effect of black mother on birth weight. According to the results, we suggest that mothers need to decrease the number of cigarettes smoked during pregnancy because it has a negative effect on the weight of the newborn baby, we also suggest the model's estimation in the future study, which includes all other relevant explanatory variables in the model which have a significant effect on birth weight.

Reference

Rubin, D., Leventhal, J., Krasilnikoff, P., Weile, B. and Berget, A., 1986. Effect of passive smoking on birth-weight. *The Lancet*, *328*(8504), pp.415-417.

Mills, J.L., Graubard, B.I., Harley, E.E., Rhoads, G.G. and Berendes, H.W., 1984. Maternal alcohol consumption and birth weight: How much drinking during pregnancy is safe?. *Jama*, *252*(14), pp.1875-1879.

Olsen, J., Rachootin, P. and Schiødt, A.V., 1983. Alcohol use, conception time, and birth weight. *Journal of Epidemiology & Community Health*, *37*(1), pp.63-65.

Wang, N., Tikellis, G., Sun, C., Pezic, A., Wang, L., Wells, J.C., Cochrane, J., Ponsonby, A.L. and Dwyer, T., 2014. The effect of maternal prenatal smoking and alcohol consumption on the placenta-to-birth weight ratio. *Placenta*, *35*(7), pp.437-441.

Appendix

Appendix A

Descriptive Statistics (Gretl Output)

Summary Statistics, using the observations 1 - 1832

(missing values were skipped)

Variable	Mean	Median	S.D.	Min	Max
fage	31.9	31.0	5.71	18.0	64.0
bwght	3.40e+003	3.43e+003	577.	360.	5.20e+003
fmaps	9.00	9.00	0.480	2.00	10.0
cigs	1.09	0.000	4.22	0.000	40.0
drink	0.0198	0.000	0.289	0.000	8.00
male	0.514	1.00	0.500	0.000	1.00
mblck	0.0595	0.000	0.237	0.000	1.00
fwhte	0.890	1.00	0.313	0.000	1.00
fblck	0.0584	0.000	0.235	0.000	1.00

Appendix B

Regression Model (1) (Gretl Output)

Model 1: OLS, using observations 1-1832 (n = 1714)

Missing or incomplete observations dropped: 118

Dependent variable: bwght

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	3422.79	14.2035	241.0	<0.0001	***
cigs	−11.9718	3.39967	−3.521	0.0004	***
drink	−8.60616	48.3620	−0.1780	0.8588	

Mean dependent var	3409.934	S.D. dependent var	571.1588
Sum squared resid	5.55e+08	S.E. of regression	569.3182
R-squared	0.007595	Adjusted R-squared	0.006435
F(2, 1711)	6.546973	P-value(F)	0.001471
Log-likelihood	−13304.93	Akaike criterion	26615.86
Schwarz criterion	26632.20	Hannan-Quinn	26621.90

Appendix C

Regression Model (2) (Gretl Output)

Model 2: OLS, using observations 1-1832 (n = 1706)

Missing or incomplete observations dropped: 126

Dependent variable: bwght

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	1036.32	268.930	3.853	0.0001	***
cigs	−11.8670	3.29073	−3.606	0.0003	***
drink	−14.9484	46.4735	−0.3217	0.7478	
fage	6.74739	2.33621	2.888	0.0039	***
fmaps	210.748	27.9459	7.541	<0.0001	***
male	95.0948	26.5429	3.583	0.0003	***
mblck	−227.722	129.559	−1.758	0.0490	***
fwhte	240.494	60.1410	3.999	<0.0001	***

fblck	428.237	139.335	3.073	0.0021	***

Mean dependent var	3412.826	S.D. dependent var	562.1306
Sum squared resid	5.06e+08	S.E. of regression	546.1077
R-squared	0.060623	Adjusted R-squared	0.056195
F(8, 1697)	13.68967	P-value(F)	2.17e-19
Log-likelihood	−13168.80	Akaike criterion	26355.60
Schwarz criterion	26404.58	Hannan-Quinn	26373.73

Appendix D

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of bwght

chi2(1) = 55.57

Prob > chi2 = 0.0000

Appendix E

Ramsey RESET test using powers of the fitted values of bwght

Ho: model has no omitted variables

F(3, 1694) = 6.76

Prob > F = 0.0002