REGRESSION ANALYSES OF THE PHILIPPINE BIRTH WEIGHT DISTRIBUTION

Elline Jade Beltrano, Robert Neil Leong
De La Salle University

ABSTRACT. Low birth weight has both short-term and long-term effects. It can lead to complications among infants causing neonatal deaths. Several literatures also suggested relationships between low birth weight and delayed mental and physical development. These negative effects are further magnified in developing countries, one of which is the Philippines. In this paper, birth weight is analysed through logistic, ordinary least squares, and quantile regression techniques using a sample from the 2008 Philippine Birth Recode. Quantile regression results offer a more dynamic picture of how these correlates affect the conditional distribution of birth weight. The obtained estimates of the marginal effects of several demographical and maternal health correlates of birth weight suggest that socially and economically impoverished mothers are more likely to have low birth weight babies. These results would recommend a focus on improving maternal health care through proper education.

Keywords: *birth weight, quantile regression, logistic regression, ordinary least squares*

1. INTRODUCTION

A particular target in the Millennium Development Goals (MDG) is to reduce child mortality by two-thirds between 1990 and 2015. One key indicator of this goal is a decreased infant mortality rate, which for the Philippines would entail a decline from 57.0 per 1,000 live births (1990) to 19.0 per 1,000 live births (2015). Though figures from 2008 (24.9 per 1,000 live births) suggested a high pace of progress, it is still probable that the Philippines will not be able to achieve its target by 2015.

Most literatures have found strong associations between infant mortality and low birth weight (LBW). Although LBW is not a direct cause, the complications due to it (e.g. inability to maintain body temperature) account for 13.8% and 15.3% of infant deaths in the Philippines for the years 2006 and 2007, respectively. Also, these complications currently rank as the third leading cause of infant deaths both locally and globally (Reolalas and Novilla, 2010).

The United Nations Children's Fund (UNICEF) defined LBW babies as newborns weighing less than 2,500 grams with the measurement taken within the first hour of life. Globally, 15.50% of total live births in 2008 are of LBW classification. In the Philippines, 21.20% of live births in 2008 are classified as LBW babies which is the largest for the past 23 years. Currently, the country ranks as the 14th (out of 225 countries) with the highest incidence of LBW cases (WHO, 2012).

Aside from significant associations with infant mortality, LBW also has other negative effects particularly on physical and mental development of children. A series of medical studies now known as the *Barker's Hypothesis* has found that reduced fetal growth is strongly associated with many chronic conditions (e.g. cardiovascular disease, diabetes, obesity) in later life (Barker, 1997). In another study, LBW children are more likely to delay entry into school or attend special classes (Corman and Chaikind, 1998) suggesting a direct link between birth weight and intelligence quotient. In the light of socio-economic concerns, LBW babies result in higher economic costs for society (e.g. higher health care costs and lower labor market payoffs). Even worse, socio-economic inequality causes great disparity between LBW outcomes (Lewit *et al.*, 1995).

For these reasons, there is a need to better understand the determinants of birth weight, both demographical and behavioral, as well as the extent of their impact. While several studies focused either only on LBW outcomes through logistic regression or on general marginal effects of several determinants through ordinary least squares regression, it would be better to study how these factors affect different conditional birth weight distributions. A particular study has focused on the above rationale which made use of quantile regression as its primary statistical technique (Abrevaya, 2001). The aforementioned study has served as the motivation for this research to be done in the local setting. There are two objectives in this study: (1) on a medical aspect, to provide a better understanding of the impact of various maternal characteristics and pregnancy behaviour on the distribution of birth weight, and (2) on a statistical aspect, to introduce quantile regression as a technique in analysing data requiring assessment of marginal effects of the covariates on the different conditional quantile distributions.

2. METHODOLOGY

2.1 Data

The data used in this study was the 2008 Philippine Birth Recode from the National Demographic and Health Survey. Data management and preparation were aided by medical literatures and by an expert obstetrician-gynecologist.

From the original data set, only 14 variables were kept which are essential to the study. These variables contain information on maternal characteristics (age of mother at birth of child, type of place of residence, region of residence, educational attainment, wealth index, total children ever born, and preceding birth interval), maternal behavior during pregnancy (smoking status, pregnancy complications, prenatal care measures, and iron supplementation), and birth outcome (birth weight in grams and gender of child). Moreover, only 3,327 cases were taken for the study using the following inclusion criteria: (1) birth weight entry is non-missing and valid (i.e. at least 500 grams), (2) child is born alive from a single birth outcome, and (3) information on prenatal measures is existent.

2.2 Analysis

Regression models fitted on the data ordinary least squares, and quantile models. For the quantile regression models, five quantiles were considered: $\tau =$ 0.05, 0.20, 0.50, 0.80, and 0.95. Full models were fitted for all quantile models while stepwise selection was applied for the logistic model. For ordinary least squares, both full and stepwise models were fitted. Diagnostics, goodness-of-fit, and accuracy tests were done afterwards. All statistical procedures and inferences were done at a significance level α of 0.05 using Stata/SE 10.

Table 5. Quantile regression and OLS regression estimates (full models)

	Quantile Regression					OLS
	5%	20%	50%	80%	95%	
AGE	57.10 (52.12)	*63.89** (29.46)	12.73 (15.28)	15.87 (22.05)	-1.73 (31.11)	26.12 (17.57)
AGESQ	-1.04 (0.87)	*-1.21** (0.49)	-0.38 (0.26)	-0.42 (0.35)	-0.13 (0.51)	-0.5855 (0.30)
PLACE	-190.69 (104.88)	-15.07 (38.48)	4.97 (27.80)	-9.06 (35.16)	-12.52 (51.03)	-8.84 (26.72)
REGION1	117.15 (196.62)	*138.23** (55.11)	*138.46** (42.22)	96.73 (70.35)	-0.85 (88.04)	*146.13** (39.66)
REGION2	*-502.77** (198.28)	-4.83 (72.37)	*90.58** (43.54)	-38.01 (60.08)	*-196.35** (86.86)	-37.62 (41.04)
REGION3	-272.57 (228.73)	38.19 (62.46)	27.32 (50.12)	-64.72 (68.42)	-143.06 (85.20)	-20.26 (46.07)
REGION4	-275.36 (184.08)	35.33 (69.85)	*101.61** (46.44)	32.36 (64.31)	59.72 (90.76)	41.38 (43.50)
EDUC1	47.99 (105.20)	*205.10** (67.80)	*115.75** (44.14)	62.13 (54.63)	-25.64 (66.55)	*90.72** (35.95)
EDUC2	25.90 (142.75)	*170.05** (74.50)	62.67 (51.35)	58.60 (62.90)	-58.77 (73.22)	76.27 (42.35)
WEALTH1	31.64 (140.76)	*128.26** (61.18)	33.52 (36.62)	21.95 (45.57)	-61.08 (64.44)	59.08 (32.50)
WEALTH2	129.37 (145.46)	*116.95** (56.80)	-0.54 (37.89)	-64.26 (45.04)	*-142.09** (66.75)	26.00 (37.71)
PARITY	16.76 (32.37)	21.96 (15.75)	*35.00** (11.58)	*45.79** (13.74)	*57.63** (26.25)	*37.59** (9.75)
BIRTHINT1	149.32 (130.83)	74.11 (62.92)	76.24 (42.29)	63.48 (56.57)	-71.42 (74.28)	62.51 (37.66)
BIRTHINT2	82.71 (145.96)	*200.80** (59.16)	*176.86** (40.09)	*122.53** (52.98)	48.48 (91.00)	*152.26** (38.23)

SMOKE	45.57 (235.24)	26.40 (90.14)	-32.04 (43.24)	-120.21 (82.26)	26.00 (131.16)	-44.52 (55.18)
VISIT	5.26 (15.06)	*12.20** (4.27)	*12.38** (4.74)	*8.82** (4.48)	0.30 (7.03)	*9.54** (3.83)
PRENATAL1	-1.38 (219.13)	-168.64 (102.42)	83.57 (63.21)	21.52 (114.29)	177.13 (166.98)	8.58 (67.12)
PRENATAL2	-49.72 (203.52)	-182.45 (102.37)	30.98 (62.16)	3.72 (110.40)	54.11 (162.55)	-30.05 (64.81)
COMPLI	-22.71 (90.89)	1.20 (50.80)	-46.64 (32.82)	-66.54 (38.22)	-57.42 (52.13)	-29.36 (30.17)
IRON	186.68 (111.97)	69.35 (64.16)	43.19 (37.18)	27.60 (48.46)	51.46 (70.43)	33.33 (41.99)
GENDER	69.08 (87.39)	60.84 (45.08)	*69.05** (24.93)	64.31 (34.34)	51.13 (45.34)	*78.94** (22.24)
constant	829.45 (764.10)	*1328.52** (445.11)	*2461.58** (215.51)	*3130.26** (338.72)	*4051.06** (436.02)	*2353.02** (256.66)

**Significant at 5% significance level*
( ) *Standard errors enclosed in parentheses*

3. RESULTS AND DISCUSSION

Table 5 shows the estimation results for the full models of both quantile regression and ordinary least squares (OLS) regression. Only a few variables are to be thoroughly discussed since they were found to behave interestingly across different quantiles.

Age of Mother at Birth of Child

AGE enters the model as a quadratic effect. The effects at various quantiles are shown in Figure 3. At the lower quantiles (τ = 0.05, 0.20), the mother's age effect tends to be more concave. It increases up to ages 26 or 27 (which are considered as the optimal ages) with magnitudes of about 780 grams and 840 grams for 5% and 20% quantiles, respectively. However, less concavity is observed for τ = 0.50 and τ = 0.80, with optimal ages of 18 to 20 years. The magnitudes of the effect are 105 grams and 150 grams for the respective quantiles. For the same quantiles, negative effects of AGE can be observed at ages 34 and 38 years, respectively. At the highest quantile (τ = 0.95), there is an immediate decline in the effect of AGE reaching optimality at the earliest age considered in the study (14 years old) with a negative magnitude of about 50 grams. In the OLS model, mother's age effect reaches optimality at around ages 22 to 24 (with magnitudes of around 290 grams) and negative effect starts at age 45. Collectively, it can be seen that for the lower quantiles, OLS seems to underestimate the quadratic effects of AGE but the opposite is observed for the other quantiles.

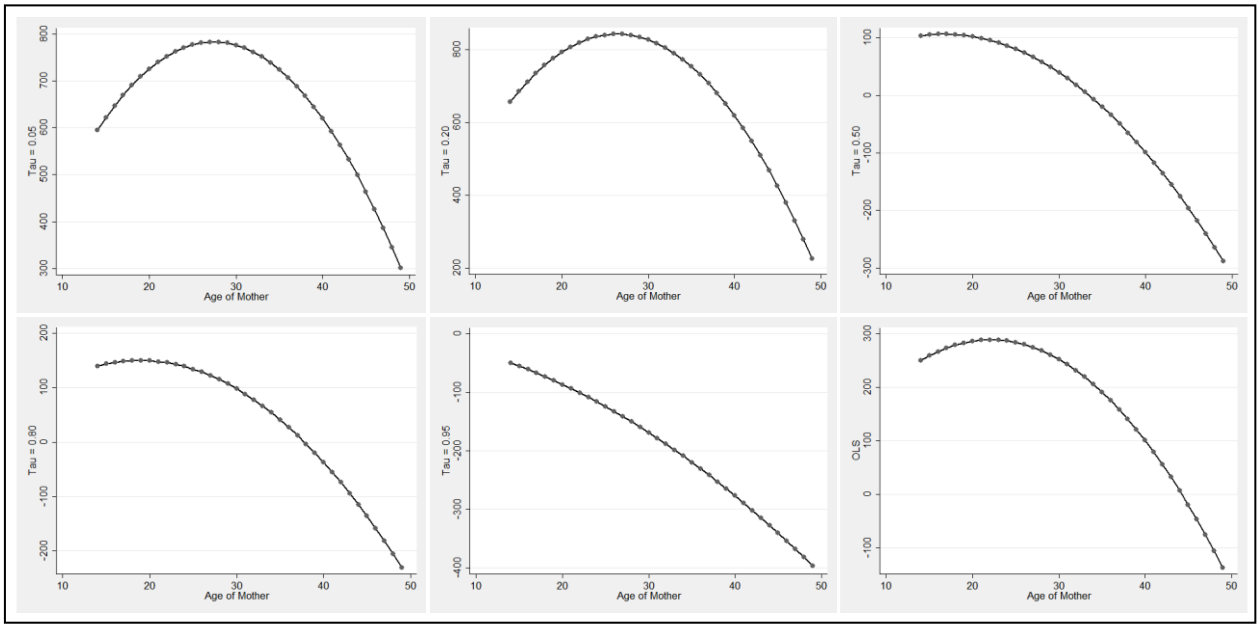

Figure 3. Quadratic effect of AGE

Educational Attainment

For the education categorical variables (EDUC1 and EDUC2), the estimates are the differences in birth weights from infants whose mothers had at most a primary education. The differentials in birth weights between mothers who attended at most a primary level education and secondary education is highest at the 20% quantile with a magnitude of about 205 grams. The disparity in birth weight is also considerably large at the median with a magnitude of almost 116 grams. For mothers who had at least a tertiary educational attainment, the birth weight differential is also highest at the 20% quantile with a magnitude of about 170 grams. These magnitudes are underestimated by OLS regression.

Wealth Index

The estimates of the wealth categorical variables are differences in birth weights from infants whose mothers belong to the poor class. The largest disparity between mothers who are in the lower class and the middle class (WEALTH1) is at the 20% quantile with a magnitude of 128 grams. This magnitude is underestimated by OLS regression. Meanwhile, the disparity between birth weights of infants born to poor mothers and infants born to rich mothers (WEALTH2) is large at different quantiles except at the median where it is very small. These large differentials seem to be underestimated by OLS regression. Also, the magnitudes of the effects of being rich tend to increase towards the extreme quantiles of the distribution. Lastly, there is a positive effect of being rich at the lower quantiles but negative effect at the upper quantiles.

From the aforementioned discussion of the marginal effects of the covariates, it should be noted that the interpretation of their causal effects may be somewhat controversial, especially across quantiles. This may be due to several unobservable factors exogenous to the study. For instance, both indicator variables for prenatal care attendants indicate negative effects at the lower quantiles suggesting that intervention by medical professionals may "decrease" birth weight. In this case, it is suspected that there is the self-selection of mothers to only seek medical help only because of the presence of pregnancy abnormalities.

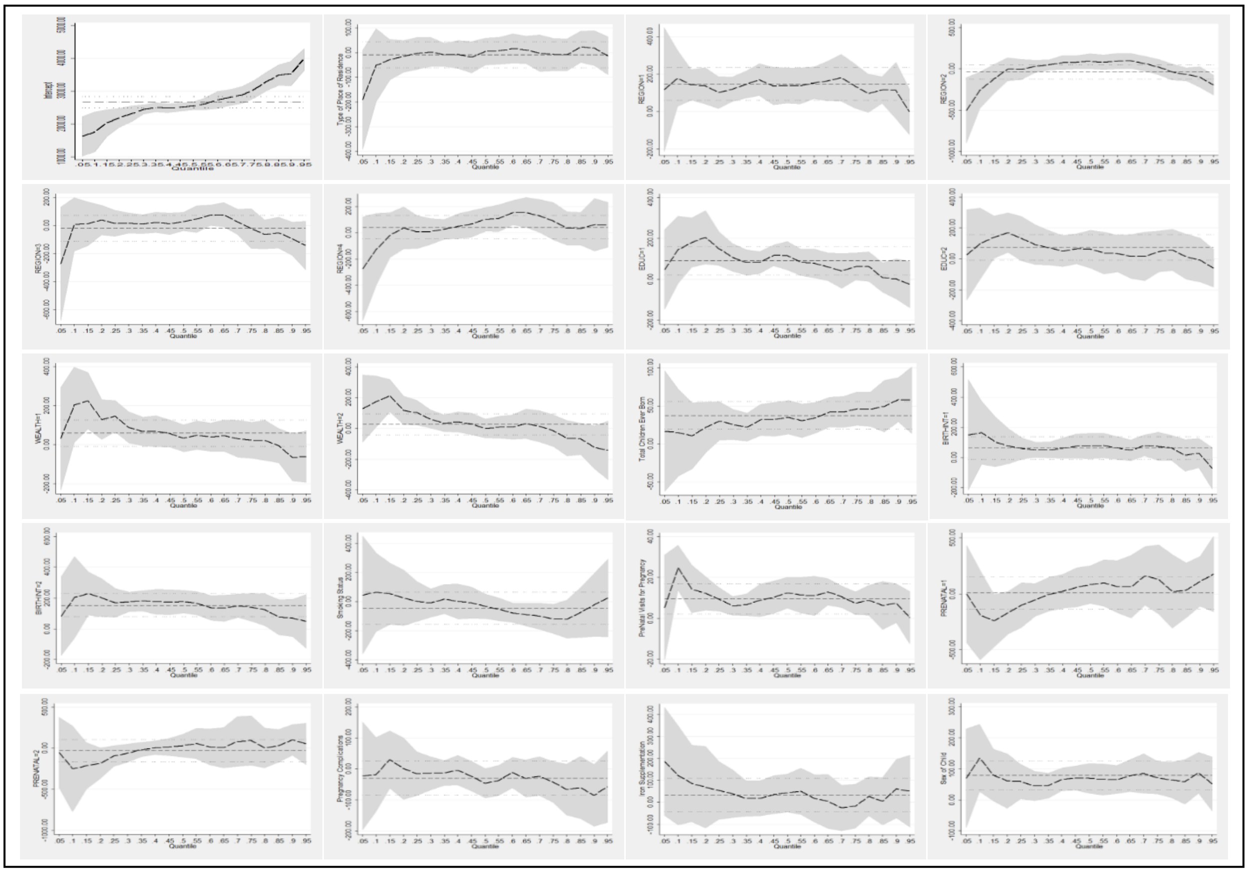

Figure 4. Marginal effects of different covariates across various quantiles

Figure 4 shows the marginal effects of different covariates across various quantiles. For several variables, the 95% confidence intervals for some quantiles seem not to overlap with their respective OLS intervals. This may suggest a significant difference between the parameter vectors $\boldsymbol{\beta}_{\boldsymbol{\tau}}$ and $\boldsymbol{\beta}_{\mathbf{OLS}}$. To formally test this, the MSS test for heteroskedasticity was used. It resulted to the rejection of the null hypothesis for quantiles 0.05, 0.20, and 0.95, with the same $p \approx 0.000$. This indicates that for the said quantiles, their parameter vector is significantly different from that of the OLS regression. From those quantiles for which their corresponding $\boldsymbol{\beta}_{\boldsymbol{\tau}}$ are unequal, the variables which caused their differences pertained to the following characteristics of the mother: age of mother at birth of child, region of residence, educational attainment, wealth index, and prenatal care attendant.

5. CONCLUSION

Quantile regression provided an interesting picture of the marginal effects of the covariates considered in the birth weight distribution. It shows fluctuation of effects of some variables across different conditional quantiles of the response variable which cannot be observed from a conditional mean model. Moreover, suggestions of relationships of variables through ordinary least squares regression may lead to negligence of certain important issues. For instance, a particular predictor may not be significant in an OLS model but it actually is at lower quantiles. Though quantile regression involves more rigorous computations, it still offers a model with fewer assumptions, thus, making it more flexible. In comparison with a logistic regression model, it allows a quantitative assessment of the marginal effects of the predictors at any distribution of the response variable rather than a likelihood association of the predictors to a certain characteristic of the response variable. However, it should be considered that in interpreting results obtained from quantile regression, proper care should be administered. That is, similar to any other regression models, coefficients should not be interpreted as causal effects due to the existence of several unobserved factors.

From the set of all variables considered in the study, several exhibited significant relationships with birth weight in all the fitted regression models: logistic, ordinary least squares, and quantile regressions. Majority of these variables are of maternal characteristics (e.g. region of residence, educational attainment, total number of children ever born, and preceding birth interval in months). From the constructed models, similar information can be extracted from logistic, conditional mean, and conditional median models. However, the same cannot be said in considering other conditional quantiles. That is, there are instances when variables behave differently across these quantiles. Specifically at the lower quantiles which represents the low birth weight infants, it is evident that socially and economically impoverished mothers are marginalized in terms of birth weight outcomes. Furthermore, most of them were also underprivileged in areas concerning maternal health. Thus, it is recommended that policy-makers should give focus on improving both the accessibility and quality of prenatal health care for mothers, especially among those who are marginalized. One way of doing this is through proper education, both on the side of mothers and public health workers.

6. ACKNOWLEDGEMENT

The authors would like to extend their sincerest and most heartfelt gratitude to Dr. Ma. Carmen Quevedo, M.D., for sharing her time and expertise with regards to the medical background of the analysis, and to Measure Demographic and Health Surveys (DHS), for providing the essential data utilized in the research.

BIBLIOGRAPHY

Abrevaya, J. (2001). The effects of demographics and maternal behavior on the distribution of birth outcomes. *Empirical Economics*, *26*, 247-257.Reolalas, A. T., & Novilla, M. M. (2010). *Newborn deaths in the Philippines*. 11th National Convention on Statistics (NCS), (pp. 6-7).

Barker, D. (1997). Maternal nutrition, fetal nutrition, and disease in later life. *Nutrition*, *3*, 807-813.

Corman, H., & Chaikind, S. (1998). The effect of low birthweight on the school performance and behavior of school-aged children. *Economics of Education Review 17*, 307-316.

Lewit, E. M., Barker, L. S., Corman, H., & Shiono, P. H. (1995). The direct costs of low birth weight. *The Future of Children 5*, 35-51.

National Statistics Office (NSO) [Philippines] and ICF Macro. 2012. *National Demographic and Health Survey 2008* [Dataset]. PHBR52.DTA. Calverton, Maryland: National Statistics Office and ICF Macro [Producers]. ICF Macro [Distributor], 2012.

World Health Organization. World health statistics 2012. Geneva: WHO Press, 2012.