1.5.3 Infant Birth Weight

The third example reconsiders an investigation by Abreveya (2001) of the impact of various demographic characteristics and maternal behavior on the birth weight of infants born in the United States. Low birth weight is known to be associated with a wide range of subsequent health problems and has even been linked to educational attainment and labor market outcomes. Consequently, there has been considerable interest in factors influencing birth weight and public policy initiatives that might prove effective in reducing the incidence of low-birth-weight infants.

Although most of the analysis of birth weight has employed conventional least-squares regression methods, it has been recognized that the resulting estimates of various effects on the conditional mean of birth weights were not necessarily indicative of the size and nature of these effects on the lower tail of the birth-weight distribution. In an effort to focus attention more directly on the lower tail, several studies have recently explored binary response (e.g., probit) models for the occurrence of low birth weights – conventionally defined to be infants weighing less than 2500 grams. Quantile regression offers a natural complement to these prior modes of analysis. A more complete picture of covariate effects can be provided by estimating a family of conditional quantile functions.

The analysis will be based on the June 1997 Detailed Natality Data published by the National Center for Health Statistics. Like Abreveya's study, the sample is restricted to singleton births, with mothers recorded as either black or white, between the ages of 18 and 45, resident in the United States. Observations with missing data for any of the variables described in the following were also dropped from the analysis. This process yielded a sample of 198,377 babies. Education of the mother is divided into four categories: less than high school, high school, some college, and college graduate. The omitted category is "less than high school," so coefficients must be interpreted relative to this category. The prenatal medical care of the mother is also divided into four categories: those with no prenatal visit, those whose first prenatal visit was in the first trimester of the pregnancy, those with the first visit in the second trimester, and those with the first visit in the last trimester. The omitted category is the group with a first visit in the first trimester; they constitute almost 85 percent of the sample. The other variables are, hopefully, self-explanatory.

Figure 1.11 presents a concise summary of the quantile regression results for this example. Each plot depicts one coefficient in the quantile regression model. The solid line with filled dots represents the point estimates, $\{\hat{\beta}_j(\tau) : j = 1, \ldots, 16\}$, with the shaded gray area depicting a 90% pointwise confidence band. Superimposed on the plot is a dashed line representing the ordinary least-squares estimate of the mean effect, with two dotted lines again representing a 90% confidence interval for this coefficient.

In the first panel of the figure, the intercept of the model may be interpreted as the estimated conditional quantile function of the birth-weight distribution of

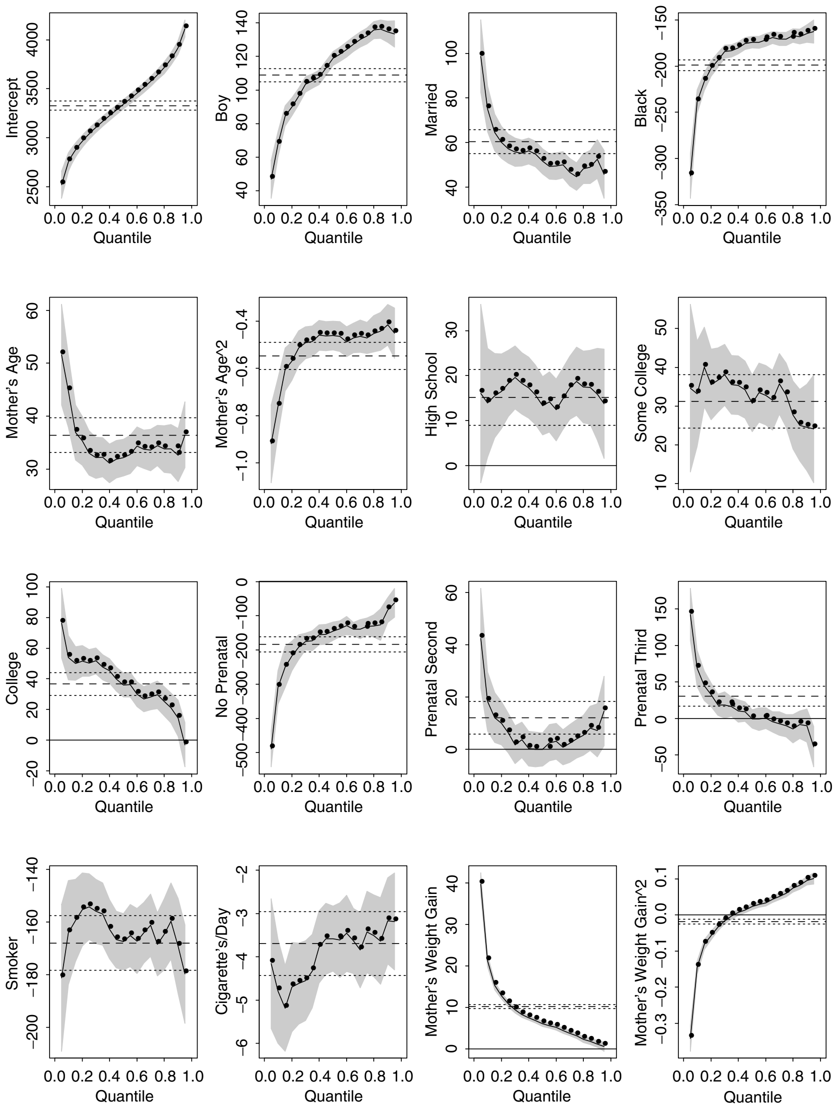

Figure 1.11. Quantile regression for birth weight.

a girl born to an unmarried, white mother with less than a high school education, who is 27 years old and had a weight gain of 30 pounds, did not smoke, and had her first prenatal visit in the first trimester of the pregnancy. The mother's age and weight gain are chosen to reflect the means of these variables in the sample. Note that the $\tau = 0.05$ quantile of this distribution is just at the margin of the conventional definition of a low-birth-weight baby.

Boys are obviously bigger than girls, about 100 grams bigger according to the ordinary least-squares (OLS) estimates of the mean effect, but, as is clear from the quantile regression results, the disparity is much smaller in the lower quantiles of the distribution and somewhat larger than 100 grams in the upper tail of the distribution. At any chosen quantile we can ask how different the corresponding weights of boys and girls are, given a specification of the other conditioning variables. The second panel answers this question.

Perhaps surprisingly, the marital status of the mother seems to be associated with a rather large positive effect on birth weight, especially in the lower tail of the distribution. The public health implications of this finding should, of course, be viewed with caution, however.

The disparity between birth weights of infants born to black and white mothers is very large, particularly at the left tail of the distribution. The difference in birth weight between a baby born to a black mother and a white mother at the 5th percentile of the conditional distribution is roughly one-third of a kilogram.

Mother's age enters the model as a quadratic. At the lower quantiles the mother's age tends to be more concave, increasing birth weight from age 18 to about age 30, but tending to decrease birth weight when the mother's age is beyond 30. At higher quantiles there is also this optimal age, but it becomes gradually older. At the third quantile it is about 36, and at $\tau = 0.9$ it is almost 40. This is illustrated in Figure 1.12.

Education beyond high school is associated with a modest increase in birth weight. High school graduation has a quite uniform effect over the whole range of the distribution of about 15 grams. This is a rare example of an effect that really does appear to exert a pure location shift effect on the conditional distribution. Some college education has a somewhat more positive effect in the lower tail than in the upper tail, varying from about 35 grams in the lower tail to 25 grams in the upper tail. A college degree has an even more substantial positive effect, but again much larger in the lower tail and declining to a negligible effect in the upper tail.

The effect of prenatal care is of obvious public health policy interest. Since individuals self-select into prenatal care, results must be interpreted with considerable caution. Those receiving no prenatal care are likely to be at risk in other dimensions as well. Nevertheless, the effects are sufficiently large to warrant considerable further investigation. Babies born to mothers who received no prenatal care were on average about 150 grams lighter than those who had a prenatal visit in the first trimester. In the lower tail of the distribution this effect is considerably larger – at the 5th percentile it is nearly half a kilogram! In contrast, mothers who delayed prenatal visits until the second or third trimester have substantially *higher* birth weights in the lower tail than mothers who had a visit in the first trimester. This might be interpreted as the self-selection effect of mothers confident about favorable outcomes. In the upper three quarters of the distribution there seems to be no significant effect.

Smoking has a clearly deleterious effect. The indicator of whether the mother smoked during the pregnancy is associated with a decrease of about 175 grams in birth weight. In addition, there is an effect of about 4 to 5 grams per cigarette per day. Thus, a mother smoking a pack per day appears to induce a birth-weight

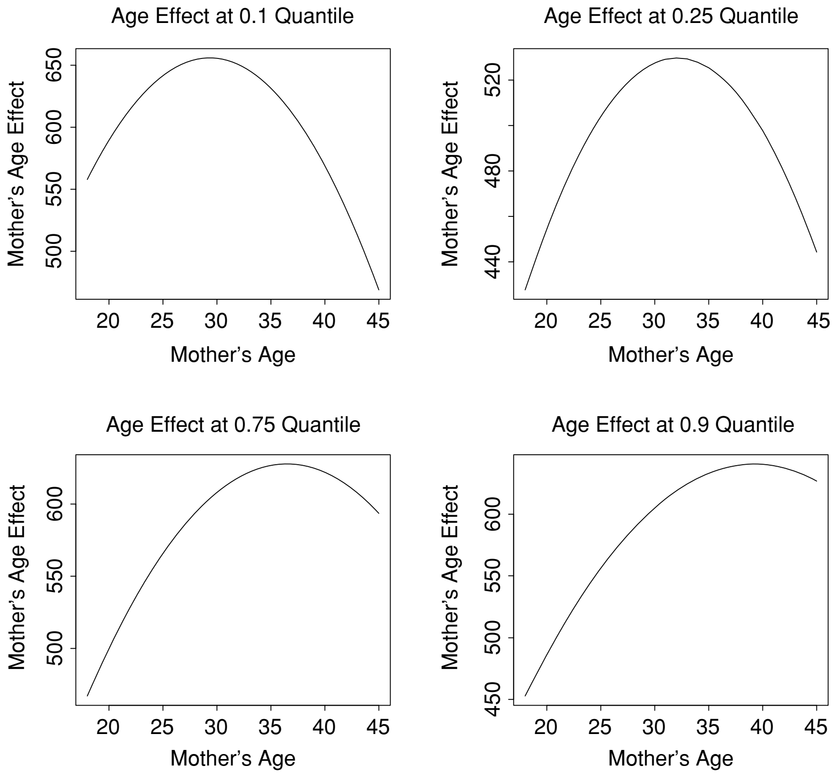

Figure 1.12. Mother's age effect on Birth weight. The estimated quadratic effect of mother's age on infant birth weight is illustrated at four different quantiles of the conditional birth-weight distribution. In the lower tail of the conditional distribution, mothers who are roughly 30 years of age have the largest children, but in the upper tail it is mothers who are 35–40 who have the largest children.

reduction of about 250 to 300 grams, or from about one-half to two-thirds of a pound. In contrast to some of the other effects, the effect of smoking is quite stable over the entire distribution, as indicated by the fact that the least-squares point estimates of the two smoking effects are (nearly) covered by the quantile regression confidence band.

Lest this smoking effect be thought to be attributable to some associated reduction in the mother's weight gain, one should hasten to point out that the weight gain effect is explicitly accounted for with a quadratic specification. Not surprisingly, the mother's weight gain has a very strong influence on birth weight, and this is reflected in the very narrow confidence band for both linear and quadratic coefficients. Figure 1.13 illustrates the marginal effect of weight gain by evaluating over the entire range of quantiles for four different levels of weight gain. At low weight gains by the mother, the marginal effect of another

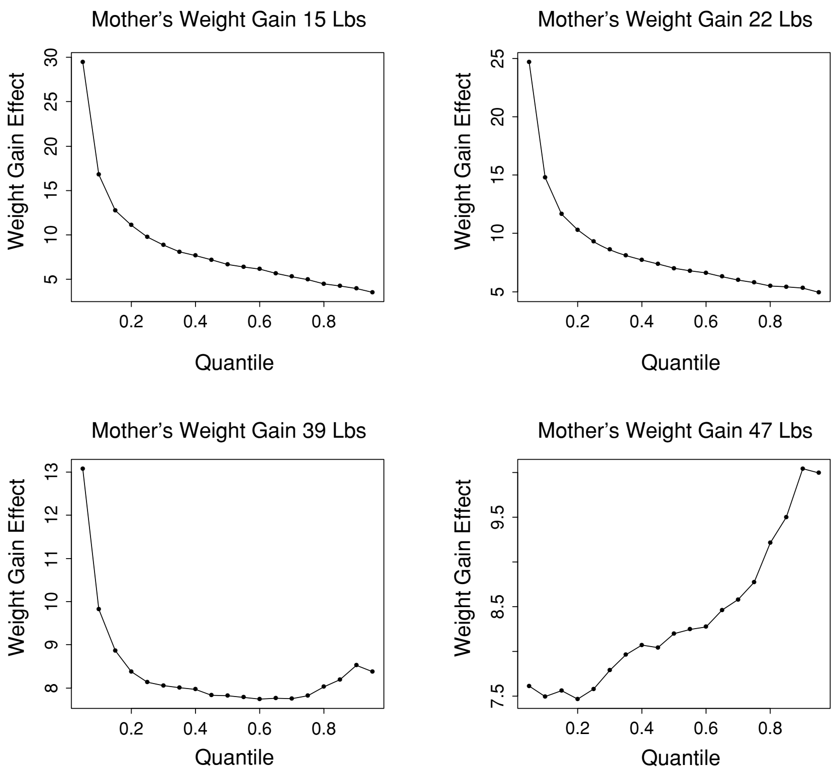

Figure 1.13. Mother's weight gain marginal effect. The marginal effect of the mother's weight gain, again parameterized as a quadratic effect, tends to decrease over the entire range of the conditional distribution of birth weight. Thus, incremental weight gain is most influential in increasing the weight of low-birth-weight infants. But for mothers with unusually large weight gains, this pattern is reversed and the effect is largest in the upper tail of the conditional birth-weight distribution.

pound gained is about 30 grams at the lowest quantiles and declines to only about 5 grams at the upper quantiles. This pattern of declining marginal effects is maintained for large weight gains, until we begin to consider extremely large weight gains, at which point the effect is reversed. For example, another pound gained by the mother who has already gained 50 pounds has only a 7-gram effect in the lower tail of the birth-weight distribution, and this increases to about 10 grams at the upper quantiles. The quadratic specification of the effect of mother's weight gain offers a striking example of how misleading the OLS estimates can be. Note that the OLS estimates strongly suggest that the effect is linear with an essentially negligible quadratic effect. However, the quantile regression estimates give a very different picture, one in which the quadratic effect of the weight gain is very significant except where it crosses the zero axis at about $\tau = 0.33$.