Journal of Management Science and Engineering 6 (2021) 336–343

Contents lists available at ScienceDirect

Journal of Management Science and Engineering

journal homepage: www.keaipublishing.com/en/journals/journal-of-management-science-and-engineering/

A new quantile treatment effect model for studying smoking effect on birth weight during Mother's pregnancy

Shengfang Tang [a], Zongwu Cai [b], Ying Fang [a, c], Ming Lin [a, c, *]

[a] Wang Yanan Institute for Studies in Economics and Fujian Key Laboratory of Statistical Sciences, Xiamen University, Xiamen, 361005, China
[b] Department of Economics, University of Kansas, Lawrence, KS, 66045, USA
[c] Department of Statistics and Data Science, School of Economics, Xiamen University, Xiamen, 361005, China

ARTICLE INFO

Article history:
Available online 19 June 2021

Keywords:
Birth weight
Heterogeneity
Quantile regression
Smoking
Treatment effect

ABSTRACT

This paper proposes a new quantile regression model to characterize the heterogeneity for distributional effects of maternal smoking during pregnancy on infant birth weight across different the mother's age. By imposing a parametric restriction on the quantile functions of the potential outcome distributions conditional on the mother's age, we estimate the quantile treatment effects of maternal smoking during pregnancy on her baby's birth weight across different age groups of mothers. The results show strongly that the quantile effects of maternal smoking on low infant birth weight are negative and substantially heterogenous across different ages.

1. Introduction

There is an increasing number of interests in estimating causal effect of a treatment, program or policy on economic or other outcomes, which is important in comparing different medical treatments, social programs and intervention policies. A conventional approach is to assume that the treatment variable is unconfounded after conditioning on a set of observable covariates, which leads to regression, matching or inverse probability weighting based methods, among others, for estimating the average treatment effects.[1] However, although the average treatment effects are interesting and critical measures in determining a treatment or policy's effect, they might not be sufficient to fully capture the impact of the treatment or policy except under very restrictive conditions. In particular, when the treatment effect is heterogeneous, the average treatment effects can not provide enough information about the effect of the treatment or policy on other points in the outcome distribution which a researcher may be more interested in. Such distributional aspects beyond simple averages can be of

* Corresponding author. Wang Yanan Institute for Studies in Economics and Fujian Key Laboratory of Statistical Sciences, Xiamen University, Xiamen, Fujian, 361005, China.
E-mail address: linming50@xmu.edu.cn (M. Lin).

[1] For example, Heckman et al. (1998), Hahn (1998), Hirano et al. (2003), Abrevaya et al. (2015), Abadie and Imbens (2006, 2016), and the references therein.

https://doi.org/10.1016/j.jmse.2021.06.005

fundamental interest and arise naturally in many areas of empirical economic research.[2] To characterize the heterogeneous effects along with the outcome distribution, quantile treatment effect (QTE), originally introduced by Doksum (1974) and Lehmann (1975), provides an effective and intuitive tool to document such heterogeneity. The importance of QTE has motivated researchers to consider its identification and estimation under common identifying assumptions such as selection on observables (see Assumption 2.1(i) later) as in Firpo (2007) and Donald and Hsu (2014). Recently, Tang (2020) provided a comprehensive survey on recent developments in modeling methods for the quantile treatment effect.

It has been explicitly recognized in program evaluation literature that the effect of a treatment can be heterogeneous across different individuals (Heckman and Robb (1985); Heckman et al. (1997)). Therefore, besides the treatment effect for the entire population, it may also be of interest to estimate the treatment effect in various covariates of interest. Some of this interest in treatment effect heterogeneity has been extensively considered in the recent literature. For example, to characterize the heterogeneous effect across different sub-populations, Abrevaya et al. (2015) and Lee et al. (2017) considered the partially conditional average treatment effect.

In this paper, we propose a parametric quantile treatment effect model to characterize the heterogeneous distributional effect of treatment variable on outcome of interest across different covariates of interest. More specifically, a polynomial restriction is proposed to specify the quantile functions of the potential outcome distributions conditional on the covariates of interest. The rationale for using a polynomial model is that the notation is simple and the estimator is computationally convenient and the resulting estimator simplifies to a linear quantile regression as in Koenker and Bassett (1978). Different from Koenker and Bassett (1978), there are missing data in the potential outcomes. Hence, the unknown parameters in the conditional quantile functions can not be estimated directly by the quantile regression estimator as in Koenker and Bassett (1978). To this end, we provide conditions (see Assumption 2.1 later) under which the conditional quantiles of the potential outcome distributions can be recovered from the distribution of the observable data.

In our empirical study, the proposed model is applied to reexamine the quantile treatment effect of a first-time mother's smoking during pregnancy on the birth weight of her baby conditional on the mother's age. Some previous work has investigated the effect of maternal smoking during pregnancy on infant birth weight; see for example, Abrevaya (2006), Abrevaya and Dahl (2008), Abrevaya et al. (2015), Lee et al. (2017), and the reference therein. In particular, Abrevaya et al. (2015) and Lee et al. (2017) investigated the average effect of mother's smoking during pregnancy on infant birth weight across different age groups of mothers. However, Cai et al. (2020) argued that the heterogeneity of the effect of mother's smoking on infant birth weight exists so that it might not be enough to consider the average treatment effect. Instead, they explored the nonparametric quantile effects of maternal smoking during pregnancy on her baby's birth weight across different age groups of mothers. In this paper, we apply the proposed parametric model to re-investigate the quantile treatment effect of mother's smoking during pregnancy on infant birth weight across the mother's age. It is worth stressing that our goal is not to provide another estimate of the quantile treatment effect of mother's smoking on the birth weight of her child per se, but rather to illustrate how the proposed parametric model can be easily used to explore the heterogeneity of the quantile treatment effect across the mother's age.

The structure of the paper is as follows. In the next section, we introduce the model framework and the estimation procedure. Section 3 presents an empirical application and Section 4 concludes the paper.

2. Econometric modeling

2.1. Framework

The focus of our interest is to estimate the causal effect of a binary treatment, represented by the binary variable D, on some scalar outcome variable, denoted by Y. More specifically, $D = 1$ denotes that some individual in a large sample receives the active treatment, while $D = 0$ denotes the same individual receives the control treatment. Following Rubin (1974) and Imbens and Rubin (2015), treatment effects are defined in terms of potential outcomes. Specifically, $Y(1)$ denotes the potential outcome for the individual under treatment and $Y(0)$ stands for the potential outcome under control. For each individual, the treatment variable D and the outcome Y are observed, which has the following relationship with the potential outcomes $Y(1)$ and $Y(0)$ as $Y = D \cdot Y(1) + (1 - D)Y(0)$. In addition, a vector of pretreatment variables or covariates, denoted by X, are observed too, which are predetermined relative to the treatment assignment and oftentimes contain characteristics of the units measured before the treatment assignment is known.

Recently, the quantile treatment effect has been popularly studied and extensively applied in the economics literature and real applications.[3] Specifically, for $j = 0$ and 1, let $F_{Y(j)}(y)$ be the distribution function of the potential outcome $Y(j)$. Then, for a given quantile level $\tau \in (0, 1)$, the QTE is defined as $F^{-1}_{Y(1)}(\tau) - F^{-1}_{Y(0)}(\tau)$, where for $j = 0$ and 1, $F^{-1}_{Y(j)}(\tau)$ is the τth quantile function of the distribution $F_{Y(j)}(y)$, given by $F^{-1}_{Y(j)}(\tau) = \inf\{y \in \mathbb{R} : F_{Y(j)}(y) \geq \tau\}$. The quantile treatment effect can be used to

[2] For example, in trying to understand the effect of a government-subsidized training program, the effect of the program on the lower tail of the savings distribution is of central interest for welfare analysis; see LaLonde (1995) and Abadie et al. (2002) for more details. Similarly, when evaluating effects of unionization on wage inequality, distributional effect is also of central interest for welfare analysis; see Freeman (1980) and Card (1996) for more discussions.

[3] See, for example, Firpo (2007), Donald and Hsu (2014), Chernozhukov and Hansen (2005, 2006), Frölich and Melly (2013), and the references therein.

characterize the heterogeneous distributional effect of a treatment or policy on outcome of interest. However, as stated in the introduction, in many applications, besides the population quantile treatment effect, researchers may also be interested in estimating the heterogeneous distributional effect of a treatment or policy on outcome of interest across different sub-populations defined by certain covariates of interest. To this end, the focus in this paper is on the quantile treatment effect conditional on a subset of the pretreatment variables, termed as the partially conditional quantile treatment effect. Specifically, let $W = (W_1, \ldots, W_k) \in \mathbb{R}^k$ be a sub-vector of $X \in \mathbb{R}^p$, where $1 \leq k < p$. Similarly, let $F_{Y(j)|W}(y\,|w)$ denote the distribution function of $Y(j)$ conditional on $W = w$ for $j = 0$ and 1. Then, the partially conditional quantile treatment effect is defined as

$$\delta_\tau(w) = Q_{1,\tau}(w) - Q_{0,\tau}(w),$$

where $Q_{j,\tau}(w)$ is the τth quantile function of the distribution of $Y(j)$ conditional on W, given by $Q_{j,\tau}(w) = \inf\{y \in \mathbb{R} : F_{Y(j)|W}(y\,|\,w) \geq \tau\}$ and $\tau \in (0, 1)$ is the quantile level. As mentioned in the introduction, to capture the heterogeneous effect across different sub-populations, Abrevaya et al. (2015) and Lee et al. (2017) considered the partially conditional average treatment effect, which is defined as

$$\alpha(w) = E(Y(1)|W = w) - E(Y(0)|W = w)$$

It is noted that the partially conditional average treatment effect can only capture the heterogeneity across different sub-populations defined by the covariate W, while the proposed partially conditional quantile treatment effect in this paper can capture the heterogeneity for both across-distribution and across different sub-populations simultaneously. Consequently, it can provide more information about the effect of treatment or policy variable on outcome of interest.

It is easy to see that estimation of quantile treatment effect is complicated by the fact that for each individual in the population, we observe at most one of the potential outcomes $Y(1)$ and $Y(0)$. Therefore, the parameter of interest $\delta_\tau(w)$ can not be identified without further restrictions. To solve the identification problem, it is assumed in this paper that assignment to treatment is unconfounded as in Rubin (1980) and Rosenbaum and Rubin (1983), and that the probability of assignment is bounded away from 0 to 1. Formally, the following assumptions are imposed.

Assumption 2.1. For almost every $x \in \mathcal{X}$, where $\mathcal{X}$ is the support of X,

(i) $(Y(0), Y(1))$ is independent of D conditional on $X = x$;
(ii) $0 < \underline{c} \leq p(x) = P(D = 1|X = x) \leq c < 1$ for some $0 < \underline{c} < c < 1$, where $p(x)$ is the so-called propensity score function.

Assumption 2.1(i) is often referred to as "unconfoundedness" or "selection on observables" and it states that when a vector of measured covariates are controlled, the treatment variable is independent of the potential outcomes. It rules out the existence of unobserved factors that affect treatment assignment and are also correlated with the potential outcomes. Assumption 2.1(ii) implies that, for almost all values of X, the population includes treated and untreated units. Clearly, it is easy to show that indeed, $\delta_\tau(w)$ can be identified under Assumption 2.1 so that Assumption 2.1 holds true through the whole paper.

2.2. Model specification and estimation procedure

In this section, our focus is on estimating $\delta_\tau(w)$ based on a linear model for the conditional quantile functions of the potential outcome distributions conditional on $W = w$. Formally, a linear restriction for the conditional quantile functions is given by

$$Q_{j,\tau}(w) = \alpha_{j,\tau} + \beta_{j,\tau}^\top w \tag{1}$$

for $j = 0$ and 1, where $\alpha_{j,\tau} \in \mathbb{R}$ and $\beta_{j,\tau} \in \mathbb{R}^k$ are the unknown parameters. Based on the specification in (1), it is easy to see that the treatment effect parameter of interest $\delta_\tau(w)$ has the form of

$$\delta_\tau(w) = (\alpha_{1,\tau} - \alpha_{0,\tau}) + (\beta_{1,\tau} - \beta_{0,\tau})^\top w.$$

Note that the particular form of the conditional quantile functions specified in (1) can be easily extended to nonlinear parametric models.

According to the approach proposed by Bassett and Koenker (1982), the parameters of the conditional quantile functions in (1) can be expressed as

$$(\alpha_{j,\tau}, \beta_{j,\tau}) = \arg\min_{(\alpha,\,\beta) \in \mathbb{R}^{k+1}} E[\rho_\tau(Y(j) - \alpha - \beta^\top W)] \tag{2}$$

for $j = 0$ and 1, where $\rho_\tau(u)$ is the check function, defined as $\rho_\tau(u) = u \cdot (\tau - I\{u < 0\})$. It is important to note that the potential outcomes $Y(1)$ and $Y(0)$ are not observed simultaneously. Therefore, the minimization problem in (2) can not be solved directly. To overcome this problem, the following weight functions of D and X are defined

$$\kappa_1(X,D) = D/p(X), \quad \text{and} \quad \kappa_0(X,D) = (1-D)/[1-p(X)].$$

Then, it follows by the unconfoundedness assumption and the law of iteration expectation that

$$\begin{aligned} &E\big[\kappa_j(X,D)\cdot\rho_\tau(Y-\alpha-\beta^\top W)\big] \\ &= E\big\{E\big[\kappa_j(X,D)\rho_\tau(Y-\alpha-\beta^\top W)\big|X,D=1\big]\cdot P(D=1|X)\big\} \\ &\qquad + E\big\{E\big[\kappa_j(X,D)\rho_\tau(Y-\alpha-\beta^\top W)\big|X,D=0\big]\cdot P(D=0|X)\big\} \\ &= E\{E[\rho_\tau(Y(j)-\alpha-\beta^\top W)|X]\} = E[\rho_\tau(Y(j)-\alpha-\beta^\top W)] \end{aligned}$$

for $j = 0$ and 1. Therefore,

$$\begin{aligned} (\alpha_{j,\tau},\beta_{j,\tau}) &= \arg\min_{(\alpha,\,\beta)\in\mathbb{R}^{k+1}} E[\rho_\tau(Y(j)-\alpha-\beta^\top W)] \\ &= \arg\min_{(\alpha,\,\beta)\in\mathbb{R}^{k+1}} E\big[\kappa_j(X,D)\cdot\rho_\tau(Y-\alpha-\beta^\top W)\big] \end{aligned} \tag{3}$$

for $j = 0$ and 1. It should be noted that the weights $\kappa_1(X, D)$ and $\kappa_0(X, D)$ are non-negative, so both the population objective function in (3) and its sample analogy are globally convex in (α, β) and hence the global optimum can be obtained in a finite number of iterations. Following the analogy principle in Manski (1988), a natural estimator of $(\alpha_{j,\tau}, \beta_{j,\tau})$ is the sample counterpart of (3).

To be specific, it is assumed that a random sample of n observations on $\{Y_i, X_i, D_i\}_{i=1}^n$ is observed. If the weight function $\kappa_j(X_i, D_i)$ for $j = 0$ and 1 would be known, then the estimation problem becomes a weighted quantile regression as discussed in Newey and Powell (1990). Unfortunately, the propensity score function $p(x)$ is unknown in general in practice, so these two weight functions are also unknown. Therefore, to estimate the unknown parameters $\alpha_{j,\tau}$ and $\beta_{j,\tau}$, one needs an estimation in two steps. First, one can use some standard parametric or nonparametric techniques to estimate the unknown function $p(x)$. Since the treatment variable D is binary, a logistic or probit regression model can be used to estimate $p(x)$ in the parametric case.[4]

Once the estimator $\hat{p}(x)$ of $p(x)$ is obtained, the fitted values $\{\hat{\kappa}_j(X_i, D_i)\}$ can be used to estimate $(\alpha_{j,\tau}, \beta_{j,\tau})$ in the second step as follows, for $j = 0$ and 1,

$$(\hat{\alpha}_{j,\tau},\hat{\beta}_{j,\tau}) = \arg\min_{(\alpha,\,\beta)\in\mathbb{R}^{k+1}} \frac{1}{n}\sum_{i=1}^{n} \hat{\kappa}_j(X_i,D_i)\,\rho_\tau(Y_i-\alpha-\beta^\top W_i). \tag{4}$$

3. Empirical analysis

In this section, the proposed parametric quantile regression model for treatment effect is applied to the analysis of the quantile treatment effect of maternal smoking during pregnancy on infant birth weight while allowing for arbitrary treatment effect heterogeneity conditional on the mother's age.

3.1. Data description

We use the same dataset as in Abrevaya et al. (2015), collected by the North Carolina State Center Health Services. Similar to Abrevaya et al. (2015), our focus is on the sample for the first-time pregnant white mothers with 433, 558 observations.

It is well documented that there are many factors to cause low infant birth weight and maternal smoking during pregnancy has been labeled the most important preventable negative cause of low infant birth weight; see Kramer (1987) for the detailed discussion. Although it is well understood that there is a negative relationship between maternal smoking during pregnancy and infant birth weight, it is not an easy task to accurately estimate the causal effect of maternal smoking during pregnancy on infant birth weight. Actually, there were some previous attempts to estimate the causal effect of maternal smoking on infant birth weight by using an instrumental variable strategy or panel data approach; see, for example, Evans and Lien (2005), Abrevaya (2006), Abrevaya and Dahl (2008), and the references therein. Recently, the program evaluation approach is successfully employed to estimate this effect; see, for example, Almond et al. (2005), Walker et al. (2009), Abrevaya et al. (2015), and Lee et al. (2017). In particular, both Abrevaya et al. (2015) and Lee et al. (2017) explored how the average effect of maternal smoking on infant birth weight changes across different age groups of mothers. They found different degrees of heterogeneity by age and the main qualitative finding is that smoking has a more severe impact at higher ages. As discussed in Abrevaya and Dahl (2008), the costs associated with birth weight have been found to exist primarily at

[4] Of course, one can use a nonparametric procedure to estimate $p(x)$, and several approaches have been proposed in the treatment effect literature. Examples include the series logit estimator (SLE) as in Hirano et al. (2003) and Firpo (2007), local polynomial regression estimation as in Ichimura and Linton (2005) and higher order kernel regression estimation as in Abrevaya et al. (2015). The reader is referred to the original references for details.

the low end of the infant birth weight distribution (with costs increasing significantly at the very low end). Therefore, researchers may be more interested in understanding the distributional effect of maternal smoking on baby's birth weight. This motivates us to consider the estimation of the quantile treatment effect of maternal smoking on infant birth weight. To be specific, our interest is to estimate how the quantile effect of smoking changes across different age groups of mothers. To this end, the proposed model is used to estimate the quantile effect of maternal smoking during pregnancy on infant birth weight conditional on the mother's age.

Since our interest is to estimate how the quantile effect of maternal smoking during pregnancy changes across different age groups of mothers, the conditional variable W is the mother's age. In addition, D denotes the treatment variable which is equal to one if the mother smokes and zero otherwise. The outcome of interest Y is the baby's birth weight measured in grams. In this example, $Y(1)$ denotes the infant birth weight for the treated (smoking) group and $Y(0)$ stands for the infant birth weight for the untreated (no-smoking) group. Fig. 1 depicts the kernel density estimations of infant birth weight for the un-treated (solid line) and treated (dotted line) groups, respectively. It looks that the density estimations of infant birth weight both for the treated and untreated groups are asymmetric and fat-tailed in the left side. To further confirm these findings, the sample skewness and kurtosis of infant birth weights are computed and also, a symmetry test is conducted to see if the distributions are symmetry. The results are reported in Table 1, which support the findings observed from Fig. 1. Based on the discussions above, it gives us a strong motivation to consider the distributional effect of maternal smoking during pregnancy on infant birth weight.

To estimate the causal effect of maternal smoking on infant birth weight using program evaluation procedure, the main difficulty is to find certain baseline covariates such that the potential birth weight outcomes are independent of the smoking decision conditional on the baseline covariates. Several studies have been conducted in the literature by using variants of the unconfoundedness assumption to identify the average effect of smoking on infant birth weight; see, for example, Almond et al. (2005), Da Veiga and Wilder (2008), and Walker et al. (2009) for details. In this paper, the same set of covariates X as in Abrevaya et al. (2015) are used, including certain parents' background information and the mothers' medical and health records as well as other variables. Specifically, Table 2 summarizes the pre-treatment covariates used in the dataset.

Another practical problem is how to estimate the unknown propensity score function $p(x)$. Following Abrevaya et al. (2015), a logit model for the propensity score function is considered and the explanatory variables used in the logit model consist of all the elements of X displayed in Table 2, the square of the mother's age, and the interaction terms between the mother's age and all other elements of X. Finally, the partially conditional quantile treatment effect, $\delta_\tau(w)$, is estimated for mother's ages in the interval from 20 to 30.

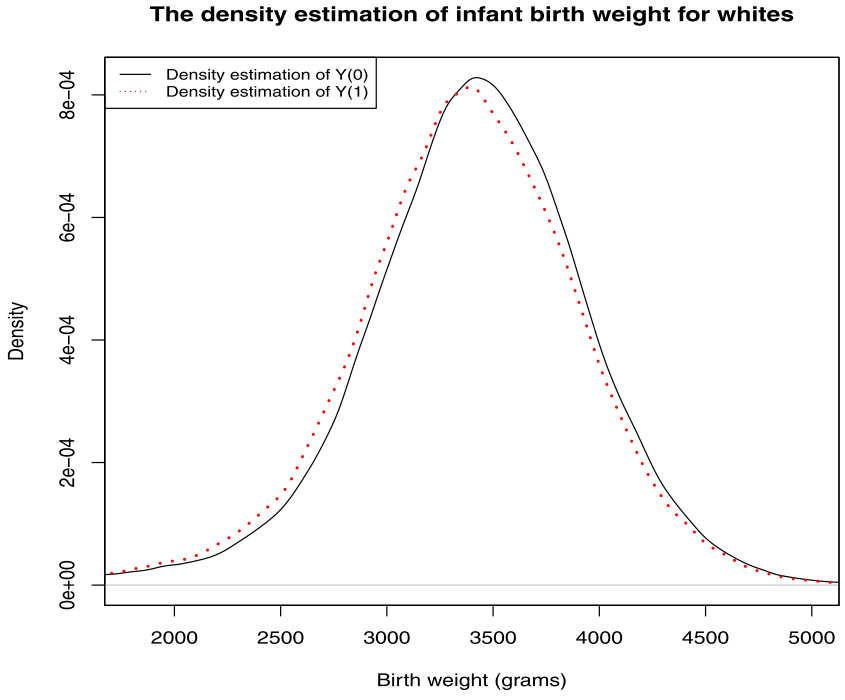

Fig. 1. The kernel density estimation of infant birth weight for white mothers. The solid line is for $Y(0)$ and the dotted line for $Y(1)$.

Table 1
Descriptive statistics and symmetry testing results.

	$Y(0)$	$Y(1)$
Mean	3398.68	3346.85
Skewness	−0.846	−0.840
Kurtosis	5.931	5.734
Symmetry test (p-value)	0.000	0.000
Number of observations	359,172	74,386

Table 2
Variable description.

		Variable name	Variable type	Description
Outcome of interest		bweight	continuous variable	Infant birth weight measured in grams
Treatment variable		smoke	binary	Whether mother smokes or not during pregnancy?
Pre-treatment covariates (X):	Certain parents' background information			
		mage	continuous variable	Mother's age
		fage	continuous variable	Father's age
		meduc	integer	Mother's years of schooling
		feduc	integer	Father's years of schooling
		fagemiss	integer	Whether or not father's age is missing?
		married	binary	Whether or not mother is married?
	Mothers' medical and health records			
		prenatal	integer	Month of first prenatal visit which is equal to 10 if prenatal care is foregone
		prenatal-visits	integer	Number of prenatal visits
		hyperpr	binary	Whether or not mother suffered from hypertension?
		terms	integer	Previous (terminated) pregnancies
		anemia	binary	Whether or not mother suffered from anemia?
		ultra	binary	Whether or not mother took ultra sound exams?
		diabetes	binary	Whether or not mother suffered from gestational diabetes?
		amnio	binary	Whether or not mother took amniocentesis?
	Other variables			
		male	binary	Whether or not baby is male?
		drink	binary	Whether or not mother used alcohol?

3.2. Empirical results

As stated above, since the costs associated with birth weight exist primarily at the low end of the birth weight distribution, we consider three quantile levels: $\tau = 0.1$, 0.25 and $\tau = 0.5$. First, a polynomial model is explored and it turns out that the following models for $Q_{0,\tau}(w)$ and $Q_{1,\tau}(w)$ are suitable

$$Q_{j,\tau}(w) = \alpha_{j,\tau} + \beta_{j,\tau}\cdot\text{age} + \gamma_{j,\tau}\cdot\text{age}^2$$

for $j = 0$ and 1. Therefore,

$$\delta_\tau(w) = (\alpha_{1,\tau} - \alpha_{0,\tau}) + (\beta_{1,\tau} - \beta_{0,\tau})\cdot\text{age} + (\gamma_{1,\tau} - \gamma_{0,\tau})\cdot\text{age}^2,$$

which is a quadratic form of age. Table 3 summarizes the estimation results from the computer output in the package *quantreg* in R.[5] It is noted that for simplicity, the standard error is computed by using the nonparametric bootstrap approach.[6] It can be seen from Table 3 that all the estimated coefficients are statistically significant at the 1% level. In particular, $\hat{\gamma}_{j,\tau}$ is significant too for three values of τ (0.1, 0.25 and 0.50). Based on these estimated coefficients, $\hat{\delta}_\tau(w)$ can be computed easily and it is

[5] The command *rq()* in the package *quantreg* in R with the weights as in (4) is employed to obtain the estimated coefficients $\hat{\alpha}_{j,\tau}$, $\hat{\beta}_{j,\tau}$, and $\hat{\gamma}_{j,\tau}$.

[6] To obtain the precise standard error, one needs to derive the accurate theoretical expression for the standard error of the estimated parameters, which is full of challenges and is beyond the scope of the present paper. Therefore, for simplicity, here we use the nonparametric bootstrap approach to calculate the standard error of the estimated coefficients.

Table 3
Estimation results of the coefficients.

Coefficients	Quantile regressions		
	10%	25%	50%
α_0	1850.455***	2414.536***	2819.315***
	(53.042)	(32.012)	(24.562)
β_0	71.734***	52.741***	46.288***
	(4.282)	(2.647)	(2.031)
γ_0	−1.378***	−0.984***	−0.845***
	(0.083)	(0.054)	(0.041)
α_1	1793.233***	2299.294***	2679.412***
	(133.221)	(80.272)	(72.806)
β_1	63.283***	50.057***	44.556***
	(11.176)	(6.547)	(6.039)
γ_1	−1.311***	−1.027***	−0.870***
	(0.224)	(0.132)	(0.120)

Note: The dependent variable is infant birth weight (in grams). *: significant at 10% level; **: 5% level; ***: 1% level. Standard errors of the estimates are reported in the parentheses.

displayed in Fig. 2, plotting the estimation results for the partially conditional quantile treatment effect, $\hat{\delta}_\tau(w)$, at three quantile levels $\tau = 0.1$, $\tau = 0.25$ and $\tau = 0.5$, respectively. For example,

$$\hat{\delta}_{0.10}(w) = -57.222 - 8.451 \cdot \text{age} + 0.067 \cdot \text{age}^2.$$

However, since $\hat{\gamma}_{1,\tau} - \hat{\gamma}_{0,\tau} = 0.067$ for $\tau = 0.10$ (-0.043 for $\tau = 0.25$ and -0.025 for $\tau = 0.50$) is too small, the linear term in the estimated curve $\hat{\delta}_\tau(w)$ dominates the whole curve in the range of (20, 30) (look like a linear), which can be seen in Fig. 2.

From Table 3 and Fig. 2, one can see clearly that the results exhibit a number of striking features. First, one can observe that there is a significant negative effect of smoking on infant birth weight across all ages and quantile levels considered. Second, the estimation results show substantial heterogeneity across different ages. For example, the estimated effect ranges from about $-200g$ to $-250g$ at the quantile level $\tau = 0.1$ as the mother's age increases from 20 to 30 and similar pattern can be observed for other quantile levels considered. Another interesting feature of the results is that for a given age, the numerical values of the quantile treatment effect point estimates at lower quantiles are bigger than those at the higher quantiles, and for a given quantile level, the estimated quantile treatment effects become stronger (more negative) at higher ages. These findings are new in the literature.

4. Conclusion

In this paper, a parametric quantile regression model for treatment effect is proposed to analyze the data from the North Carolina State Center Health Services, which have previously been used by Abrevaya et al. (2015), Lee et al. (2017) and Cai et al.

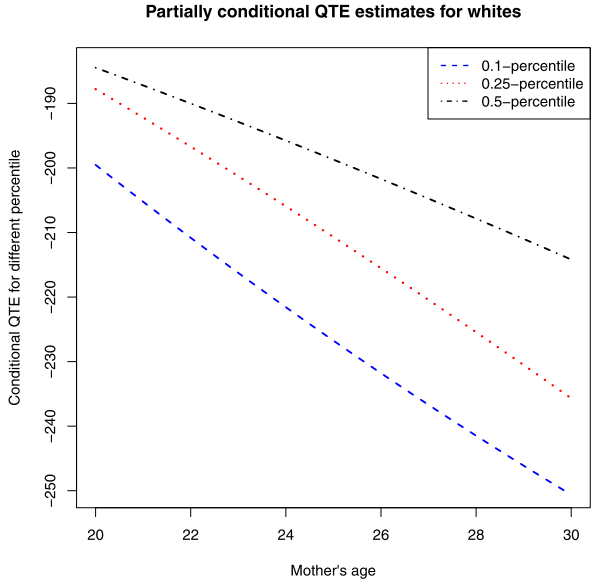

Fig. 2. Estimation results for the partially conditional quantile treatment effects.

(2020), to examine the quantile effect of maternal smoking during pregnancy on infant birth weight across different age groups of mothers. The estimated results suggest that the quantile effect of maternal smoking on infant birth weight is heterogeneous across different age groups of mothers. The results also indicate that for a given quantile level, the effect is monotonically decreasing in the mother's age and for a given mother's age, the numerical values of the effect at lower quantiles are bigger than those at the higher quantiles. These results demonstrate that estimates of treatment effect which focus on a single feature of the outcome distribution, for example, the average treatment effect, may fail to capture the full impact of the treatment variable on the outcome of interest and that examining additional features may provide more information about the relationships between the treatment variable and the outcome of interest.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank the Guest Editor, Professor Yingfei Qi and two anonymous referees for the valuable and help comments and suggestions that improved significantly the quality of the paper. Also, they acknowledge the partially financial supports from the National Natural Science Foundation of China (NSFC) for Distinguished Scholars (71625001), the NSFC key projects with grant numbers 71631004, 72033008 and 71131008, and Science Foundation of Ministry of Education of China (19YJA910003).

References

Abadie, A., Angrist, J., & Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica, 70*(1), 91–117.
Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica, 74*(1), 235–267.
Abadie, A., & Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica, 84*(2), 781–807.
Abrevaya, J. (2006). Estimating the effect of smoking on birth outcomes using a matched panel data approach. *J. Appl. Econom., 21*(4), 489–519.
Abrevaya, J., & Dahl, C. M. (2008). The effects of birth inputs on birthweight: evidence from quantile estimation on panel data. *J. Bus. Econ. Stat., 26*(4), 379–397.
Abrevaya, J., Hsu, Y.-C., & Lieli, R. P. (2015). Estimating conditional average treatment effects. *J. Bus. Econ. Stat., 33*(4), 485–505.
Almond, D., Chay, K. Y., & Lee, D. S. (2005). The costs of low birth weight. *Q. J. Econ., 120*(3), 1031–1083.
Bassett, G., & Koenker, R. (1982). An empirical quantile function for linear models with iid errors. *J. Am. Stat. Assoc., 77*(378), 407–415.
Cai, Z., Fang, Y., Lin, M., & Tang, S. (2020). *Estimating Partially Conditional Quantile Treatment Effects. Working Paper*. Department of Economics, University of Kansas.
Card, D. (1996). The effect of unions on the structure of wages: a longitudinal analysis. *Econometrica, 64*(4), 957–979.
Chernozhukov, V., & Hansen, C. (2005). An iv model of quantile treatment effects. *Econometrica, 73*(1), 245–261.
Chernozhukov, V., & Hansen, C. (2006). Instrumental quantile regression inference for structural and treatment effect models. *J. Econom., 132*(2), 491–525.
Da Veiga, P. V., & Wilder, R. P. (2008). Maternal smoking during pregnancy and birthweight: a propensity score matching approach. *Matern. Child Health J., 12*(2), 194–203.
Doksum, K. (1974). Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *Ann. Stat., 2*(2), 267–277.
Donald, S. G., & Hsu, Y.-C. (2014). Estimation and inference for distribution functions and quantile functions in treatment effect models. *J. Econom., 178*(3), 383–397.
Evans, W. N., & Lien, D. S. (2005). The benefits of prenatal care: evidence from the pat bus strike. *J. Econom., 125*(1–2), 207–239.
Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica, 75*(1), 259–276.
Freeman, R. B. (1980). Unionism and the dispersion of wages. *Ind. Labor Relat. Rev., 34*(1), 3–23.
Frölich, M., & Melly, B. (2013). Unconditional quantile treatment effects under endogeneity. *J. Bus. Econ. Stat., 31*(3), 346–357.
Hahn, J. (1998). On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica, 66*(2), 315–331.
Heckman, J. J., Ichimura, H., & Todd, P. E. (1998). Matching as an econometric evaluation estimator. *Rev. Econ. Stud., 65*(2), 261–294.
Heckman, J. J., & Robb, R. (1985). Alternative methods for evaluating the impact of interventions: an overview. *J. Econom., 30*(1–2), 239–267.
Heckman, J. J., Smith, J., & Clements, N. (1997). Making the most out of programme evaluations and social experiments: accounting for heterogeneity in programme impacts. *Rev. Econ. Stud., 64*(4), 487–535.
Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica, 71*(4), 1161–1189.
Ichimura, H., & Linton, O. (2005). *Asymptotic Expansions for Some Semiparametric Program Evaluation Estimators* (pp. 149–170). Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.
Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.
Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*(1), 33–50.
Kramer, M. S. (1987). Intrauterine growth and gestational duration determinants. *Pediatrics, 80*(4), 502–511.
LaLonde, R. J. (1995). The promise of public sector-sponsored training programs. *J. Econ. Perspect., 9*(2), 149–168.
Lee, S., Okui, R., & Whang, Y.-J. (2017). Doubly robust uniform confidence band for the conditional average treatment effect function. *J. Appl. Econom., 32*(7), 1207–1225.
Lehmann, E. L. (1975). *Nonparametrics : Statistical Methods Based on Ranks*. San Francisco: Holden-Day.
Manski, C. F. (1988). *Analog Estimation Methods in Econometrics*. New York: Chapman and Hall.
Newey, W. K., & Powell, J. L. (1990). Efficient estimation of linear and type i censored regression models under conditional quantile restrictions. *Econom. Theor., 6*(3), 295–317.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55.
Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol., 66*(5), 688–701.
Rubin, D. B. (1980). Randomization analysis of experimental data: the Fisher randomization test comment. *J. Am. Stat. Assoc., 75*(371), 591–593.
Tang, S. (2020). Recent developments in modeling quantile treatment effects. *Appl. Math. J. Chin. Univ., 35*(2), 220–243.
Walker, M. B., Tekin, E., & Wallace, S. (2009). Teen smoking and birth outcomes. *South. Econ. J., 75*(3), 892–907.