CREATES

Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2008-20

Headlights on tobacco road to low birthweight outcomes Evidence from a battery of quantile regression estimators and a heterogeneous panel

Stefan Holst Bache, Christian M. Dahl and Johannes Tang Kristensen

School of Economics and Management
University of Aarhus
Building 1322, DK-8000 Aarhus C
Denmark

Headlights on tobacco road to low birthweight outcomes*

Evidence from a battery of quantile regression estimators and a heterogeneous panel

Stefan Holst Bache · **Christian Møller Dahl** ·
Johannes Tang Kristensen

Abstract Low birthweight outcomes are associated with considerable social and economic costs, and therefore the possible determinants of low birthweight are of great interest. One such determinant which has received considerable attention is maternal smoking. From an economic perspective this is in part due to the possibility that smoking habits can be influenced through policy conduct. It is widely believed that maternal smoking reduces birthweight; however, the crucial difficulty in estimating such effects is the unobserved heterogeneity among mothers and the fact that estimation of conditional mean effects seems potentially inappropriate. We provide a unified view on the estimation of relationships between prenatal smoking and birthweight outcomes with quantile regression approaches for panel data and emphasize their differences. This paper contributes to the literature in three ways: i) we focus not only on one technique, but provide evidence from several approaches and highlight a variety of statistical issues; ii) the performance of the methods are thoroughly tested in a simulated environment, and recommendations are given on their appropriate use; iii) our results are based on a detailed data set, which includes many relevant control variables for socio-economic, wealth and personal characteristics.

Keywords Quantile regression · low birthweight · panel data · unobserved heterogeneity
JEL C13 · C23 · I10

* This version: April 2011. First version: May 2008. This paper was previously titled "Determinants of Birthweight Outcomes: Quantile Regression Based on Panel Data"

S. H. Bache, J. T. Kristensen
Aarhus University
School of Economics and Management and CREATES
Tel.: +45-8924-1602
E-mail: sbache@creates.au.dk
E-mail: jtang@creates.au.dk

C. M. Dahl (corresponding author)
University of Southern Denmark, Odense
Department of Economics and Management
Tel.: +45-6550-3267
E-mail: cmd@sam.sdu.dk

1 Introduction and Motivation

The potential adverse health consequences of low birthweight outcomes, along with the considerable economic burden they are believed to impose on society, have attracted much attention by researchers in both medical and economic literature. The use of birthweight as a proxy for the general health condition of infants is commonplace, as it has been linked to a vast array of health related complications, both short- and long-term.

The most severe event, perinatal mortality, has been found to be more likely in the event of a low birthweight outcome. Several studies find statistical evidence of this linkage, see e.g. Bernstein et al (2000), Almond et al (2005), and Black et al (2007). Furthermore, it is believed that low birthweight may lead to complications such as epilepsy, mental retardation, blindness, and deafness. For a review and references, see Hack et al (1995). While many of these complications are directly observable, some studies also consider less obvious socio-economic implications of low birthweight, a very popular topic being school performance. Kirkegaard et al (2006) find a graded relationship between birthweight and school performance. In a follow-up study with 5,319 Danish children aged 9–11, they conclude that the risk of reading, spelling and arithmetic disabilities is greater with low birthweight children. Similarly, Corman and Chaikind (1998) find that repeating a grade, or special class attendance, is more likely among low birthweight children. This may suggest that even future earnings and labour market outcomes may be affected by birthweight. According to Black et al (2007), this is indeed the case.

The strong evidence that low birthweight has adverse effects has naturally led to substantial efforts towards identifying the determinants of these undesirable outcomes. One such determinant which has received much attention in the literature is maternal smoking habits during pregnancy. Statistical efforts suggest a strong correlation between low birthweight and maternal smoking, see e.g. Bernstein et al (1978) and Permutt and Hebel (1989). Other studies examine the effect of smoking on some of the above-mentioned complications directly, e.g. Wisborg et al (2000), who find that smoking increases the risk of the sudden infant death syndrome, Wisborg et al (2001), who find an increased risk of still birth and infant mortality from maternal smoking, and Linnet et al (2006), who link hyperactive-distractible behaviour in preschool children to intrauterine exposure to tobacco smoke. Medical research gives several reasons why cigarette smoking may affect birthweight. An explanation that seems to stand out is that the foetus may suffer from chronic hypoxic stress as a consequence of smoking. DiFranza et al (2004) and Hofhuisi et al (2003) explain this phenomenon in part by a lowered maternal uterine blood flow and a reduction in oxygen diffusion across the placenta. An interesting observation is that smoking does not seem to have a significant adverse effect on all birth outcomes. Wang et al (2002) conclude that the association between maternal cigarette smoking and reduced birthweight is modified by maternal genetic susceptibility, after having considered two specific gene polymorphisms.

From an economic perspective, interest lies not with the individual as such, but rather with society as a whole. Maternal smoking habits are thus an especially interesting determinant since it is believed to be modifiable through policy conduct, e.g.

by regulating taxes on tobacco products or by introducing smoking prohibitions in public areas. While medical research gives much attention to why smoking causes low birthweight, the above has led economists to focus primarily on the extent of this effect, and the associated costs. This perspective has the advantage of allowing analysts to disregard the specific medical links between maternal smoking and low birthweight, when using appropriate methods.

In an attempt to estimate the direct costs associated with low birthweight, Almond et al (2005) use data from hospitals in New York and New Jersey to find that the costs peak at $150,000 (in year 2000 dollars) for newborns weighing 800 grams. In contrast, an infant weighing 2,000 grams has an estimated associated cost of $15,000. The soaring costs at the low end of the birthweight distribution highlight an important point. Using traditional mean regression will only uncover effects on the birthweight mean, i.e. infants weighing around 3,500 grams. One way to overcome this problem is to use a quantile regression approach, which can provide estimation results across the entire distribution. This is done by Abrevaya (2001) and Koenker and Hallock (2001), who find justification for the quantile approach since regression estimates vary throughout the distribution. It is, however, troublesome to consider the estimated effects as causal, because the analyses do not account for unobserved heterogeneity. Not only is the susceptibility of smoking effects among mothers different, as noted above, but there are undoubtedly many other individual characteristics which cannot be accounted for.

Econometric panel data models allow controlling for (time invariant) unobserved individual heterogeneity. However, their extension to a quantile regression framework is still somewhat limited. In a recent paper, Abrevaya and Dahl (2008) consider the extension of the "correlated random effects" model by Chamberlain (1984) to a quantile regression framework, and estimate the effects of various birth inputs on birthweight, using data from Arizona and Washington. Their results indicate that the negative effects of smoking, albeit present, are significantly lower in magnitude across all quantiles than the corresponding cross-sectional estimates.

This paper in part extends the results of Abrevaya and Dahl, using Danish data, which in itself is novel: no previous study has applied such techniques to data with this origin. The advantage of our data, relative to those used in existing literature, lies in the richness and availability of variables. Quite naturally, however, there are fewer observations due to a small geographical area. Finally, the Danish Civil Registration System allows perfect linkage of the data. Based on the idea of the above mentioned study, we consider a new correlated random effects specification for quantile regression which, at the cost of a more restricted specification, allows for the use of an unbalanced dataset and benefits from a more parsimonious amount of regressors. Finally, we consider fixed effects approaches to quantile regression, in particular we examine a model specification by Koenker (2004) and suggest a simple a two-stage fixed effects approach.

Before we delve into our birthweight application in Section 3, we take a tour in the realm of quantile regression for panel data. Our treatment offers a unified framework in which the different approaches can be discussed and compared appropriately. We are not aware of a similar discussion, and believe it to be novel and relevant more generally. Simulations will serve to illustrate features and test performance of the discussed estimation methods.

2 Econometric setup

2.1 Panel data and quantile regression - a prologue

A chief difficulty in examining the causal effect of prenatal smoking, and other relevant observable variables, on birthweight outcomes is the possible existence of influential but unobservable determinants. The identification and measurement of all such determinants is an impossible task, and it can thus be necessary to control for such unobserved effects. When repeated measurements for each individual are available, analysts will try to utilize this panel structure of the data to either filter out, or in some other way deal with e.g. time-invariant unobserved characteristics. There is a very well developed machinery with a variety of estimation procedures available for linear least-squares models, and hence the issues are here easily mitigated. Often, and indeed in the present analysis, the conditional mean of the response is not of primary interest, but rather our attention is directed towards the conditional quantiles. Not surprisingly, combining the power of panel data methods with that of quantile regression methods is not a topic of little interest. While there are methods available to do so, the topic is still relatively undeveloped, and there are some very important subtleties that often do not receive sufficient attention. In particular, one needs to be very careful in defining the quantities or parameters of interest for reasons that we will try to make clear. The main purpose of this section is to discuss some relevant procedures for panel data quantile regression. Simulations will serve both to evaluate the appropriateness of the methods and to investigate their performance. The discussion is of broader relevance than to the current analysis and we hope it will help others in choosing the best approach for their particular application.

There are many ways of introducing quantile regression. One can take a structural approach, assume a data-generating process (DGP), and describe how the error term may change the coefficients at various quantiles. It is often hard to specifically link a DPG to its quantile function. Therefore, another common approach is to think of an approximation to the quantile function directly, and not emphasize how data is generated. Angrist et al (2006) show that, in the linear case, the quantile regression estimator in a certain sense gives the best linear approximation to the true conditional quantile function. Related is the Skorohod representation where the response variable is generated by a (quantile) function that depends, amongst other things, on a rank-variable (or quantile index). We shall start the discussion with the following definition of the conditional quantile function. Let Y denote the response variable, X be a vector of covariates on which we condition, and $\tau \in (0,1)$ be the quantile index. We define

$$\mathrm{Q}_Y(\tau\,|\,X) \equiv \inf\left\{y : F_Y(y\,|\,X) \geq \tau\right\}. \tag{1}$$

Then, in Skorohod representation,

$$Y = Q_Y(U\,|\,X), \quad U\,|\,X \sim \text{uniform(0,1)}. \tag{2}$$

It is well-known that the function is a solution to the minimization problem

$$\mathrm{Q}_Y(\tau\,|\,X) \in \underset{q_\tau(X)}{\arg\min}\, \mathrm{E}\left[\rho_\tau(Y - q_\tau(X))\right], \tag{3}$$

where $\rho_\tau(u) = (\tau - 1\{u < 0\})u$, and the minimization is over all measurable functions. It is also a solution to the estimating function $\mathrm{E}[Y \leq q_\tau(X)] = \tau$, or equivalently $\mathrm{E}[\tau - 1\{Y \leq q_\tau(X)\}] = 0$. For an alternative to (3) for solving a parameterized version of the latter estimating equation, see Bache (2010).

Our discussion is focussed on linear-in-parameters subsets of functions over which (3) is minimized, i.e. $q(X;\beta(\tau)) = X^{\mathrm{T}}\beta(\tau)$, and we shall investigate some possible ways of incorporating information from repeated measurements to alleviate identification issues that arise due to characteristics not included in the model. The methods we discuss can be categorized into a fixed-effects and a correlated-random-effects framework. The terminology is carried over from their least-squares analogies, and we will not philosophize about the appropriateness of it. We will, however, now argue why these two branches distinguish themselves even more from each other in a quantile regression setting.

Consider the following much simplified setup. We imagine a population of individuals, where each can be either of two types, say $c = 0$ or $c = 1$. Types are attributes in the sense that they are constant and not under the individuals' control (e.g. one could think about genetic traits). The analyst has no information about types, i.e. for all purposes they are unobservable.

The question of interest is how a treatment x (e.g. smoking during pregnancy) affects the distribution of an outcome variable (e.g. birthweight). Figure 1 shows three distributions that may be of interest.

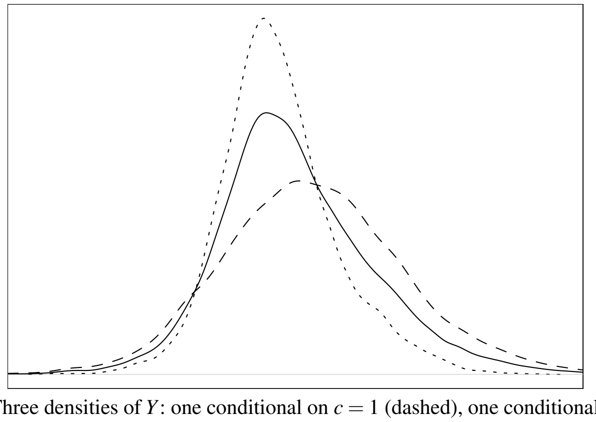

Fig. 1: Three densities of Y: one conditional on $c = 1$ (dashed), one conditional on $c = 0$ (dotted), and one unconditional of c (solid). All densities are here unconditional of x.

Now, if type is considered a part of the "noise" or "error term" in the model[1], which it may well be, say to a politician who wants to start a campaign against pre-

[1] In a treatment of quantile regression, "unexplained ranking mechanism" may be better terminology.

natal smoking, then the unconditional density is the relevant one. On the other hand, even though information in the data is not sufficient to reveal type, a doctor for example may have "inside" information on type, and would want a model that conditions on c. In this case, the conditional distributions are of interest. The main point here is that the marginal effects of changes in X on the quantile function conditioning both X and c are *not* the same as if conditioning only on X.

Fixed-effects approaches often incorporate estimates of c as a way of conditioning on it, e.g. as the estimator by Koenker (2004). We will also suggest a simple 2-stage plug-in method as a computationally simpler alternative. An inherent issue with the fixed-effects methods is the incidental parameter problem that arises when the number of repeated measurements is small and fixed.

The correlated random effects model, suggested first for regression quantiles by Abrevaya and Dahl (2008), henceforth the AD model, tries to control for dependence between X and c, which can bias the estimates, if ignored, when a "random assignment" interpretation is wanted. The idea is that one can generate one or more "sufficient covariates" from the repeated observations which carry information that can correct for the bias. For the present application, we will also consider an alternative specification that can be seen as a restricted version of the AD model, but which will allow us to use an unbalanced panel, and which is more parsimonious in terms of number of regressors.

We formally define the four estimators in section 2.3, but first we present an illustrative example of the point raised in this section about conditioning on unobserved time-invariant variables.

2.2 An illustrative simulation experiment

Consider again the simple setup with a two-type population. The model that generated the data depicted in Figure 1 is generated as

$$y_{mb} = x_{mb} + c_m + (1 + x_{mb} + c_m)\varepsilon_{mb}, \quad b = 1,2;\ m = 1,\ldots,M, \tag{4}$$

where one half of the population has $c_m = 1$, the other $c_m = 0$. The variable x_{mb} is a binary treatment and equals 1 with probability $0.5 + 0.2c_m$. The disturbance ε_{mb} is a standard normal random variable. The subscripts m and b denote "mother" and "birth" respectively, and are chosen in the light of the birthweight application.[2] Note that the unobserved type affects both location and scale of the response distribution. We also define the counterfactual variables $\tilde{x}_{mb}$ and $\tilde{y}_{mb}$, where $\tilde{x}_{mb}$ equals 1 with probability 0.5 and $\tilde{y}_{mb}$ is determined by (4) with $\tilde{x}_{mb}$ in place of x_{mb}. These then represent a counterfactual world, where type has no impact on treatment probability. We will consider three "targets" for the estimate of a coefficient for the effect of being

[2] In our application birth parity is what defines the waves, and does not represent time as such. In the examples, however, one can think of b as representing time.

treated (setting $x = 1$):

$$\Delta_\tau := Q_Y(\tau \mid x_{mb} = 1) - Q_Y(\tau \mid x_{mb} = 0) \quad (5)$$

$$\widetilde{\Delta}_\tau := Q_{\widetilde{Y}}(\tau \mid \tilde{x}_{mb} = 1) - Q_{\widetilde{Y}}(\tau \mid \tilde{x}_{mb} = 0) \quad (6)$$

$$\Delta_{\tau|c} := 1 + Q_\varepsilon(\tau) = 1 + Q_N(\tau), \quad (7)$$

where $Q_N(\tau)$ is the τth quantile of a standard normal random variable.

Let $\widehat{\beta}(\tau)$ be the estimate from a quantile regression of y_{mb} on x_{mb} and $\tilde{\beta}(\tau)$ the one from a quantile regression of y_{mb} on x_{mb} and $\bar{x}_{m\circ}$, where the latter variable is an average over the b dimension. Further, let $\check{\beta}(\tau)$ be the "oracle" estimate, where knowledge of type is assumed, from a regression of y_{mb} on x_{mb} and c_m. Table 1 shows the results from a simulation experiment, comparing these estimators with the three targets defined above for a single quantile index, $\tau = 1/5$.

	$\widehat{\beta}(0.2)$	$\widetilde{\beta}(0.2)$	$\check{\beta}(0.2)$
Δ_τ	**−0.0040** **(0.1188)**	−0.0228 (0.1619)	−0.0385 (0.1295)
$\widetilde{\Delta}_\tau$	0.0338 (0.1234)	**0.0070** **(0.1604)**	−0.0087 (0.1239)
$\Delta_{\tau\|c}$	0.0452 (0.1270)	0.0184 (0.1613)	**0.0026** **(0.1236)**

(a) $M = 999$

	$\widehat{\beta}(0.2)$	$\widetilde{\beta}(0.2)$	$\check{\beta}(0.2)$
Δ_τ	**0.0010** **(0.0380)**	−0.0289 (0.0597)	−0.0404 (0.0567)
$\widetilde{\Delta}_\tau$	0.0308 (0.0489)	**0.0009** **(0.0522)**	−0.0106 (0.0411)
$\Delta_{\tau\|c}$	0.0422 (0.0567)	0.0123 (0.0536)	**0.0008** **(0.0398)**

(b) $M = 9{,}999$

Table 1: Bias and root mean squared error (in parentheses) for the three estimators $\widehat{\beta}(\tau)$, $\widetilde{\beta}(\tau)$, and $\check{\beta}(\tau)$ against the three targets Δ_τ, $\widetilde{\Delta}_\tau$, and $\Delta_{\tau|c}$. The simulation setup is $\tau = 1/5$, "MC iterations"= 999, number of waves $B = 2$, and number of individuals $M = 999$ (panel a) and $M = 9{,}999$ (panel b). It should be noted that $\Delta_{\tau|c}$ can be calculated analytically, whereas $\widetilde{\Delta}_\tau$ and Δ_τ themselves are obtained by simulation.

The results in this example show quite clearly that the estimators estimate different quantities. A very noteworthy observation is that the estimator $\widetilde{\beta}(\tau)$, a "correlated random effects" type estimator that we will define below, does *not* try to estimate c_m, and does not suffer from an incidental parameter problem. Instead, it uses information constructed from all observations for each individual to correct for correlation between c_m and x_{mb} to get a "random assignment" interpretation. The fixed effects estimator, on the other hand, relies on some kind of estimate of c_m (above it is simply assumed known), and will most likely not perform as well as we have just seen when the number of waves is small. The last estimator is not really of interest, but it shows what happens with the usual quantile regression estimator when the treatment is endogenous.

2.3 The estimators defined

Throughout, we take Y_{mb} to be the random response variable and X_{mb} to be the corresponding covariate vector. We denote by $Z_{mb} \subset X_{mb}$ a subset that has time-varying and possibly endogenous covariates. The subscripts m and b ("mother" and "birth") are chosen to reflect the dimensions in our birthweight study. Lower case letters will denote outcomes in the sample with $m = 1, \dots, M$ and $b = 1, \dots, B_m$. The pairs $\{Y_{mb}, X_{mb}\}$ are assumed to be independent and identically distributed (IID), and so whenever no confusion arises subscripts m and b are omitted from notation.

For the purpose of this presentation of the models, we take the view that we can approximate the conditional quantile function reasonably well by a linear-in-parameters specification, and will not argue a specific DGP.[3] The following assumption states the standard QR problem in terms of the framework we shall work with to describe the panel data approaches.

*Assumption (*A.QR*): Linear quantile approximation representation. The class of functions over which* (3) *is minimized is linear, such that*

$$q_\tau(X) = q(X,\tau) = X^{\mathrm{T}}\beta(\tau) \quad \textit{and} \tag{8}$$

$$Y = q_\tau(X,U), \tag{9}$$

where U is a rank variable with $U|X \sim$ uniform$(0,1)$ *(independently of X). For any two possible ranks u_1 and u_2 it is assumed that*

$$u_1 < u_2 \Leftrightarrow q(X,u_1) < q(X,u_2) \Leftrightarrow Y_1 < Y_2. \tag{10}$$

Example: Normal location-scale model. Let

$$q(\tau,X) = X^{\mathrm{T}}\beta(\tau) = X^{\mathrm{T}}(\beta + \gamma\Phi^{-1}(\tau)), \tag{11}$$

where Φ is the standard-normal CDF*, then we have the implied familiar location-scale* DGP

$$Y = X^{\mathrm{T}}\beta + (X^{\mathrm{T}}\gamma)\Phi^{-1}(U). \tag{12}$$

The above representation in (8)–(9) is often referred to as the Skorohod representation. If there are unobserved effects that affect both X and the ranking U, then we cannot represent the problem as above and use $\beta(\tau)$ as the quantity of interest, as illustrated in the example in the previous section. In the following, we will see how one can possibly get around this problem if such unobserved characteristics are time-invariant (or here, "birth-invariant").

[3] In our simulations, of course, we need to assume some data-generating mechanism.

2.3.1 Correlated random effects quantile regression

This presentation takes a slightly different approach than the one taken in Abrevaya and Dahl (2008), but the message remains the same. We now extend the model to include unobserved characteristics C_m that partly determine Y_{mb} either directly, through $Z_{mb} \subset X_{mb}$, or both. These characteristics are assumed to be time-invariant characteristics, so dependence with Z_{mb} is one-way. The random data pairs $\{C_m, Y_{m1}, X_{m1}, Y_{m2}, X_{m2}, \dots\}$ are assumed to be IID. We wish to consider C_m a part of the unexplained ranking mechanism in the model, but at the same time control for endogenous effects propagated through Z_{mb}. To achieve this goal, we assume that repeated measurements of Z_{mb} allow for construction of sufficient covariate(s) S_m, and let the conditional quantile function of interest be $\mathrm{Q}_{Y_{mb}}(\tau \mid X_{mb}, S_m)$. Sufficiency is to be understood as to allow for the following extension of (A.QR). Again, we will often omit subscripts to simplify notation.

*Assumption (*A.CRE*): Correlated random effects representation. Consider a representation similar that of (*A.QR*), with*

$$q(X,S,\tau) = X^{\mathrm{T}}\beta(\tau) + S^{\mathrm{T}}\pi(\tau) \quad \text{and} \tag{13}$$

$$Y = q(X,S,U), \tag{14}$$

for some variable S, constructable from repeated observations of Z, such that $U|X,S \sim \text{uniform}(0,1)$.

This allows us to think of a response process $\widetilde{Y}(U) \equiv Y - S^{\mathrm{T}}\pi(U)$ as being Y "corrected" at level U for effects of C through Z, and as having τth conditional quantiles $X^{\mathrm{T}}\beta(\tau)$ for $U = \tau$. Put this way, it is emphasized that the correction gives $\beta(\tau)$ the interpretation of marginal effects in a "counterfactual world" where X is not determined by C, but where C is allowed to work directly on Y through the ranking U. Figure 2 shows a simple graph of the assumed causal relations.

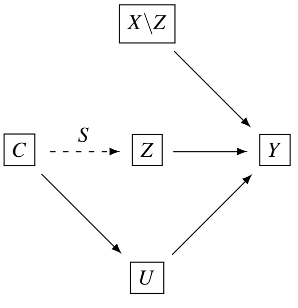

Fig. 2: The assumed relationship between variables in the model. The dashed line indicates that $Z \mid S, Y \perp U$.

Under the assumptions in (A.CRE), the quantity of interest, $\beta(\tau)$, is identified from the data, and can be estimated by means of standard quantile regression of Y on

X and S, with empirical criterion function

$$\left(\widehat{\pi}(\tau), \widehat{\beta}(\tau)\right) = \arg\min_{\pi,\beta} \sum_{m=1}^{M} \sum_{b=1}^{B_m} \rho_\tau(y_{mb} - s_m^{\mathrm{T}}\pi - x_{mb}^{\mathrm{T}}\beta). \tag{15}$$

The separability assumption that the effects from C through Z can be captured by S in a linear fashion may not be as restrictive as one would think at first, since it is allowed to vary with τ. It may therefore provide a good linear approximation at each quantile, and seems no more restrictive than believing the linear specification of q in the first-place. In general, note that the assumptions do not restrict the effect from C to a location-shift.

Example: Recall our previous introductory example. There we had a single endogenous binary treatment variable x_{mb} $(= z_{mb})$ and unobserved types c_m that affect both the probability of treatment, as well as location and scale of the outcome. Intuitively, the mean of x (over b) carries information about type, and we let $s_m = \bar{x}_{m\circ} \equiv \sum_b x_{mb}/B_m$. The simulations above confirmed that this was an acceptable choice and that the procedure was appropriate in this case.

More specifically, we mention here the following two specific CRE models, in terms of constructing S; the one proposed by Abrevaya and Dahl (2008), and another which can be seen as a special case of the first.

The "AD" model: Assume (A.CRE), that the data constitutes a balanced panel $B_m = B$, and that $S_m = (Z_{m1}^{\mathrm{T}}, \ldots, Z_{mB}^{\mathrm{T}})^{\mathrm{T}}$ are sufficient covariates.

The "CREM" model: Assume (A.CRE), and that $S_m = \bar{Z}_{m\circ}$, i.e. b-means of observed outcomes of Z_{mb}, are sufficient covariates.

The second, where the "M" in the acronym stands for *mean*, is the one deployed in the previous example. It can be seen as a special case of the AD model, in which S is essentially a weighed average, in the sense that the effect from C through Z is assumed to be the same for each b. In addition to being more parsimonious, in terms of number of regressors, this restriction allows the use of an unbalanced panel. It is possible to extend the AD model, using dummy variables, to allow for some kinds of unbalanced panels (Fitzenberger et al 2010).[4]

For the linear least-square analogues of the AD and CREM models, by Chamberlain (1984) and Mundlak (1978) respectively, one has a linear DGP, $y_{mb} = x_{mb}^{\mathrm{T}}\beta + c_m + \varepsilon_{mb}$, and one thinks of projecting c_m onto observables as

$$c_m = x_{m1}^{\mathrm{T}}\pi_1 + \cdots + x_{mB}^{\mathrm{T}}\pi_B, + \eta_m \quad \text{respectively} \tag{16}$$

$$c_m = \bar{x}_{m\circ}^{\mathrm{T}}\pi + \eta_m, \tag{17}$$

for two models. For quantile regression, using a DGP as a starting point seems restrictive as there is not necessarily a well-defined link between a DGP and a linear quantile representation. For further detail on—and a slightly different presentation of—the CRE approach we refer to the aforementioned article by Abrevaya and Dahl.

[4] We thank the editor in chief for pointing this out to us. Our results did not seem to depend on the choice of S, so we did not explore this option explicitly.

2.3.2 Fixed effects estimators

These estimators, in general, condition on the unobserved time-invariant characteristics C using some kind of estimate of them. Loosely speaking, the approaches in a way model $\mathrm{Q}_Y(\tau\,|\,X,C)$ by $\mathrm{Q}_Y(\tau\,|\,X,\widehat{C})$. As we saw in the previous simulation example, if one knows C, this approach is obviously perfectly valid, yet for a target quantity slightly different from that of the CRE approaches. Therefore, if one assumes that good estimates of C are available, these can be used in following representation.

*Assumption (*A.FE*): Quantile regression conditional on fixed effects. The class of functions over which* (3) *is minimized is linear in X and C such that*

$$q(X,C,\tau) = X^{\mathrm{T}}\beta(\tau) + C\delta(\tau) \quad \textit{and} \tag{18}$$

$$Y = q(X,C,U), \tag{19}$$

where $U|X,C \sim \text{uniform}(0,1)$ *and satisfies an ordering property equivalent to that in* (10) *of (*A.QR*).*

Essentially, the assumption here is that C is the only source of endogeneity and that it enters linearly on equal grounds with X. Conditioning on C will give $\beta(\tau)$ the same interpretation as if C was just another covariate and not a part of the unexplained ranking U.

However, in many cases the number of waves, B_m, is small (fixed) while M is large. In such a setting, good estimates of M individual effects are hard to come by, a situation referred to as the "incidental parameter problem". The estimates of C are usually not of particular interest, but it is not straight-forward to infer the consequences of M-inconsistent estimates of C on the estimates of $\beta(\tau)$.

Koenker (2004) suggested an interesting method to overcome some of the difficulties of the FE approach, in which estimation of C is intrinsic (i.e. it is a one-step estimator). The idea is to (i) estimate the model for several τs simultaneously, restricting the effect of C at each τ to be the same; and (ii) penalize estimates of the fixed effects to shrink them towards zero. Both (i) and (ii) in some sense reduce the dimensionality added by introducing estimation of fixed effects. Again, our presentation of the model is slightly alternative, compared to its original, such that it fits well into our framework.

*"*KFE*(k): Koenker's* FE QR *model". Define M parameters,* $\alpha_m = C_m\delta(\tau)$ *(for all* τ*). Assume that*

$$q(X_{mb},C_m,\tau) = X_{mb}^{\mathrm{T}}\beta(\tau) + \alpha_m \quad \textit{and} \tag{20}$$

$$Y = q(X_{mb},C_m,U_{mb}), \tag{21}$$

Let $\tau_1,\ldots,\tau_k$ *be k distinct quantile indices. Further define* $w_1,\ldots,w_k$ *to be weights that define the relative impact of each of these indices on estimation. The parameters* *in* (20) *are estimated from the following empirical criterion function:*

$$\Big(\widehat{\beta}(\tau_1),\ldots,\widehat{\beta}(\tau_k),\widehat{\alpha}_1,\ldots,\widehat{\alpha}_M\Big) = \underset{\beta_1,\ldots,\beta_k,\alpha_1,\ldots,\alpha_M}{\arg\min} \sum_{j=1}^{k}\sum_{m=1}^{M}\sum_{b=1}^{B} w_j \rho_{\tau_j}(y_{mb} - x_{mb}^{\mathrm{T}}\beta_j - \alpha_m) + \lambda \sum_{m=1}^{M} |\alpha_m|. \quad (22)$$

Let $M \to \infty$, $B \to \infty$ and $M^a/B \to 0$ for some $a > 0$. Then under some regularity conditions, $\widehat{\beta}(\tau_j)$, $j = 1,\ldots,k$, are consistent and converge to Gaussian random vectors. See (Koenker 2004, Theorem 1) for the details.

It is pretty clear from this representation that the price one pays for the gained sparseness is that unobserved fixed effects are only allowed to affect Y by location shifts. Further, often B is fixed, and consistency is not guarantied by the theorem. The parameter λ is a tuning/calibration parameter that allows one to control the impact of the penalty, and how to choose it optimally is an open research question. When $\lambda \to 0$, one has a (weighted) dummy variable regression, and when $\lambda \to \infty$ the penalty sets all FE terms to zero, effectively leaving a pure (weighted) cross-section regression. As with other penalty methods, one might consider a "post-penalty" estimation of the model where terms shrunken to zero are omitted, since non-zero terms have been affected by the penalty. On the other hand, the penalty also helps "controlling" the many FE terms by not letting them take unreasonably large values.

A simpler approach, which does not restrict C to have the same impact across quantiles in return for sparseness, and which does not need calibration, is the following two-step fixed effects estimator estimator, which resembles recent work by Arulampalam et al (2007).

"2SFE: 2-step FE QR model." Assume (A.FE). In a first step, obtain estimates $\tilde{C}_m$ of C_m from a least-squares within groups estimation. In the second step, $\tilde{C}_m$ are used in place of C in (18)*, from which estimates of $\delta(\tau)$ and $\beta(\tau)$ are obtained.*

An important implicit assumption here is that a linear approximation is appropriate for both the conditional expectation and the conditional quantiles. To hope for any asymptotic justification, one would need $B \to \infty$, which we do not consider here, as it is not relevant for our application. It has come to our knowledge that Canay (2010) has work on the asymptotic aspect of this estimator. For our purpose, performance is compared to Koenker's method and the CRE estimators in the following simulation section.

2.4 Additional simulation evidence

Indeed, the number of estimation methods under consideration allows for many interesting simulation setups. We have no desire of letting a simulation section eclipse the main part of this section or the application in question. Therefore, we have chosen a simple setup that highlights some important features and issues of the procedures discussed. In short, our findings are:

- The CRE methods do not suffer from an incidental parameters problem and perform well even when omitted effects have scale effects on the response.
- The CREM and AD models have very similar levels of performance.[5]
- FE methods generally have difficulty for short panels, and the size of the bias appears to depend critically on both τ and the actual setup. However:
  - estimating more quantiles simultaneously can be advantageous, and
  - post-estimating the "selected model" where penalized FE terms are removed can improve performance in some cases.
- The FE methods perform worse when c_m has a scale effect (i.e. effects varying over quantiles). For the KFE model, this is not surpising given its specification.
- KFE estimates are generally bounded by cross-sectional estimates and a dummy-variable quantile regression, as the theory suggests.
- Being a FE estimator, 2SFE is also cursed by incidental parameters. It benefits solely from simple estimation and from the fact that it needs no calibration.

Our simulation setup, which easily encompasses all of the estimation procedures, has the following data generating mechanism:

$$\begin{aligned} y_{mb} &= 10 - 3x_{1,mb} + x_{2,mb} + c_m + \eta_{mb} \\ \eta_{mb} &= (1 + x_{1,mb} + \gamma c_m)\varepsilon_{mb} \\ \varepsilon_{mb} &\sim N(0,1) \end{aligned} \tag{23}$$

where c_m equals 0 with probability $1/2$ and is distributed as standard normal otherwise; the binary variable $x_{1,mb}$ equals 1 with a probability, p_m, that depends on c_m in the following way:

$$p_m = \begin{cases} 0.25 & \text{if } c_m > 0.2 \\ 0.75 & \text{if } c_m < -0.2 \\ 0.50 & \text{otherwise.} \end{cases} \tag{24}$$

The idea is that the probability of treatment is only affected if individuals distinguish themselves sufficiently. We let $x_{2,mb}$ be a sum of five uniform variables on $(-0.5, 0.5)$, which then lies in $(-2.5, 2.5)$. We can think of this setup as a simplified (and scaled) simulation of our birthweight application where we have an overall intercept, from which some members of the population distance themselves (when $c_m \neq 0$). Smoking, $x_{1,mb}$, has a negative direct effect, and it has a scale effect. It is correlated with c_m which makes it necessary to control for if we want to have some notion of random assignment interpretation. The variable x_2 plays the part of "other" observables in the model. The parameter $\gamma \in \{0,1\}$ lets us control whether c_m has a scale effect in addition to its location effect, which would cause the effect from c_m to vary across quantiles.

As previously discussed, we have target quantities for the coefficient estimates on $x_{1,mb}$ that differ for CRE and FE approaches. Estimators in the latter category have the target $\beta_{fe}(\tau) = Q_N(\tau) - 3$, while those in the former have $\beta_{cre}(\tau) = Q_{c+(2+\gamma c)\varepsilon}(\tau) - Q_{c+(1+\gamma c)\varepsilon}(\tau)$.

[5] We have tried specifications where the dependence between $x_{1,mb}$ and c_m depends on b to accommodate AD. This, however, had little effect and CREM performed equally well.

	$\tau = 1/4$	$\tau = 1/2$	$\tau = 3/4$
FE	-3.6745	-3	-2.3255
CRE, $\gamma = 0$	-3.6299	-3	-2.3701
CRE, $\gamma = 1$	-3.6114	-3.0253	-2.3797

Table 2: Coefficient targets for the estimators.

We report results for the following estimators: (i) A pure dummy variable quantile regression, (ii) KFE(1) and KFE(3); (iii) a post-penalty estimation of the latter, where penalized FE terms are removed; (iv) 2SFE; and (v) the two CRE methods. We also consider a cross-sectional estimate against both targets to evaluate the bias when ignoring c_m completely. All results can be read in Tables 8 and 9 in Appendix A. Here, in the main text, we present a few selected results in Table 3. To illustrate the asymmetric performance, particularly of the FE estimators, we consider $\tau \in \{1/4, 1/2, 3/4\}$. The corresponding targets are presented in Table 2.

		$B = 2$		$B = 3$		$B = 5$	
	τ	$M = 499$	$M = 999$	$M = 499$	$M = 999$	$M = 499$	$M = 999$
Cross-section FE target	0.25	−0.3380	−0.3402	−0.3412	−0.3443	−0.3435	−0.3405
		(0.3713)	(0.3561)	(0.3618)	(0.3538)	(0.3560)	(0.3467)
	0.50	−0.3724	−0.3740	−0.3755	−0.3718	−0.3718	−0.3722
		(0.3968)	(0.3867)	(0.3920)	(0.3808)	(0.3813)	(0.3770)
	0.75	−0.4321	−0.4344	−0.4310	−0.4282	−0.4186	−0.4245
		(0.4598)	(0.4468)	(0.4489)	(0.4373)	(0.4301)	(0.4303)
Dummy regression	0.25	0.6717	0.6626	0.3531	0.3474	0.1438	0.1431
		(0.6980)	(0.6770)	(0.3813)	(0.3619)	(0.1763)	(0.1595)
	0.50	−0.0028	−0.0119	−0.0047	−0.0055	−0.0035	−0.0053
		(0.1898)	(0.1395)	(0.1272)	(0.0904)	(0.0865)	(0.0645)
	0.75	−0.6773	−0.6864	−0.3596	−0.3598	−0.1438	−0.1475
		(0.7033)	(0.7003)	(0.3877)	(0.3736)	(0.1752)	(0.1645)
KFE3 Post est.	0.25	0.0877	0.0807	0.1105	0.1081	0.0614	0.0600
		(0.1967)	(0.1425)	(0.1682)	(0.1421)	(0.1157)	(0.0905)
	0.50	−0.0224	−0.0321	−0.0474	−0.0430	−0.0253	−0.0269
		(0.1821)	(0.1357)	(0.1264)	(0.0951)	(0.0875)	(0.0676)
	0.75	−0.2531	−0.2582	−0.2052	−0.2002	−0.1040	−0.1086
		(0.3030)	(0.2844)	(0.2476)	(0.2207)	(0.1431)	(0.1279)
2SFE	0.25	0.3921	0.3844	0.2598	0.2551	0.1563	0.1553
		(0.4285)	(0.4014)	(0.2882)	(0.2707)	(0.1844)	(0.1687)
	0.50	−0.0020	−0.0110	−0.0208	−0.0210	−0.0259	−0.0267
		(0.1640)	(0.1169)	(0.1162)	(0.0844)	(0.0879)	(0.0664)
	0.75	−0.4036	−0.4080	−0.3004	−0.2981	−0.1961	−0.2005
		(0.4355)	(0.4249)	(0.3262)	(0.3101)	(0.2155)	(0.2104)
CREM	0.25	−0.0164	−0.0253	−0.0225	−0.0247	−0.0210	−0.0185
		(0.2063)	(0.1416)	(0.1439)	(0.1029)	(0.1020)	(0.0739)
	0.50	−0.0004	−0.0092	−0.0103	−0.0073	−0.0062	−0.0071
		(0.1749)	(0.1237)	(0.1290)	(0.0916)	(0.0904)	(0.0638)
	0.75	−0.0018	−0.0052	0.0048	0.0057	0.0144	0.0093
		(0.2017)	(0.1398)	(0.1431)	(0.1030)	(0.1045)	(0.0745)

Table 3: Bias and root mean squared error (rmse) for simulation of (23) with $\gamma = 0$, i.e. no scale effect of the individual effects.

A few things should be mentioned here about the results. First, the choice of penalty parameter λ in practice is an unresolved problem. For this simulation we have chosen one that approximately sets 30% of the FE terms to zero, i.e. "some but not quite enough". To get the same relative impact, this needs to be calibrated as B varies for a given N, but not vice versa. Second, the reported root mean squared error results (rmse) cannot be compared across FE and CRE methods, since targets and variance of the "total error" terms are different.

The cross section quantile regression suffers a serious bias away from both targets. Note that it does converge to some quantity in the sense that, as the sample grows, rmse and bias are more or less equal. The dummy variable regression performs very well for the median but not the other quartiles. Penalizing the FE terms and estimating all quantiles simultaneously offers improvements in the tails and post-estimating the "selected model" seems to add a little to this improvement. As expected from the theory, as B grows these models perform better. The 2SFE model performs poorly, yet better than the dummy regression, both for $\gamma = 0$ and $\gamma = 1$. If one specifies Koenker's FE method well, 2SFE is outperformed by it.

The CRE methods perform very well, and only the total sample size seems to matter (i.e. no incidental parameters curse). They seem to be close to unbiased, with diminishing variance. Table 9 in the Appendix shows that the FE methods cannot handle a scale effect of c_m, and that they become even more biased. This is not the case for the CRE methods, which still have high performance.

In conclusion, pooled cross-sectional estimation is rarely a good idea when there are omitted individual effects correlated with included variables if a "random assignment" interpretation is intended. It also appears to be the case that one can do much better than a pure dummy regression (at least in the tails). The FE methods suffer from the incidental parameters curse and high sensitivity to the data generating mechanism. It is possible to improve estimates in a FE setting by calibrating Koenker's approach, but it is hard in practice to determine whether it is done optimally.

The CRE methods in general have good performance, and do not seem to have trouble with either small B or scale effects of the individual effects. We need to stress again that they estimate something different from the FE methods, and these findings do not imply that one should discard the latter. It all depends on what the target of interest is.

In any case, with a short panel it appears that the CRE results are more reliable. From an economic policy perspective the CRE target is perhaps also more sensible in our birthweight application, as the politician is interested not in the individual as such, but in society as a whole. Why then condition on individual effects? It is part of the unexplained ranking mechanism. In this light, we will put more emphasis on the CRE results in our application. However, we will report a selection of FE results as well.

3 Data Description

We now return from our methodological excursion to put our birthweight application back in the spotlight. The data which are used throughout the analyses are in part obtained from Aarhus University Hospital, Skejby, in Denmark. In the Aarhus region this hospital is the only one with a maternity ward, and thus the data in fact represent a broad population group, i.e. all economic and social classes. Furthermore, the data are enriched with socio-economic characteristics of the mothers. These additional data have been made available by Statistics Denmark and are linked by means of the Danish Civil Registration System.

The methods we have discussed above require a panel of mothers with two or more registered births. Only singleton births are included since multiple births babies (e.g. twins) tend to be lighter. Moreover, stillbirths are excluded, and thus the population of interest are singleton live births. For the variables of interest the data offer an unbalanced panel consisting of 16,602 births and 7,900 mothers. Except for the AD model, the estimation strategies discussed allow for an unbalanced panel, and this will be the primary data set. However, in the interest of comparison with this model, estimations have also been conducted on the basis of a balanced subset of the data which have been constructed with the first two births by each mother. This includes a total of 12,670 births. The descriptive statistics for the (unbalanced) dataset are given in Table 5. The sample ranges from the year 1992 to 2005. The included variables and their role in the analysis are the topic of the remainder of this section.

The choice of birthweight as the dependent variable gives rise to an important question: should gestational age be included as an explanatory variable? There is no doubt that gestational age is correlated with birthweight. However, in the present analysis our interest lies in the total effect of maternal smoking on birthweight, including any effects propagated through gestational age. Therefore it is not necessary to include gestational age as an explanatory variable, it might even be inappropriate as it could have undesirable effects due to multicollinearity. In this context it should also be emphasised that on the matter of not including gestational age as an explanatory variable, we follow recent leading econometric studies on birthweight, see e.g. Abrevaya (2006, 2001), Abrevaya and Dahl (2008), Chernozhukov (2010), and Koenker and Hallock (2001).

The primary regressors of interest regard the mothers' smoking habits. These are represented by two separate variables: whether or not mothers smoked at the time they became pregnant (*smoked before*), and whether or not they smoked during the pregnancy (*smoked during*). Both variables are binary, as the data unfortunately do not offer details on smoked quantities. The analysis therefore cannot account for the size of the treatment, which of course may be a drawback, since it seems reasonable to believe that quantity could be important. For identification of the separate effects of the two smoke variables, it is necessary that there are mothers who actually start or stop smoking when becoming pregnant. 2,019 of 16,602 births are given by mothers who stop smoking at the time of pregnancy (12.66%). On the other hand only 10 births are given by mothers who start smoking at the time of pregnancy (0.06%). Thus the change of behavior is therefore largely one-way.

The included variables can be roughly categorised into six categories. First, there are variables relating to the behaviour of the mothers. Already mentioned are the two smoke variables. Further, we include a variable, *drink*, which indicate whether or not the mother has consumed alcohol during the pregnancy.[6] Drinking habits are also

[6] Here, alcohol is defined as consumption of more than 1 Danish standard drink (12 grams of pure alcohol) per week.

believed to be harmful to the foetus and warnings are often explicitly printed on alcoholic containers. Related to the behavioural category is also the extent to which the mother actively tried to become pregnant. Another possibility is that the pregnancy was either unwanted or accidental. To control for this, we include a dummy variable for use of *birth control pills* within four months before becoming pregnant.

A second category that seems obviously related to the health of the baby, and thus possibly birthweight, is the general health or physical ability of the mother. An important variable in this category is occurrence of pregnancy *complications*, which we represent with an aggregated dummy variable which covers things such as premature contractions, bleedings, excessive vomiting, infections, and intrauterine growth restriction. Especially this last example is important, and may be caused by factors such as high blood pressure, heart disease, malnutrition, and substance abuse. A potential problem is that tobacco smoke may also be the source of this complication, and in effect leave us with an issue of separability of effects. This will be discussed further in the next section, where we present our results. The remaining variables in this category are the number of *doctor visits* and *prenatal visits* during pregnancy, whether the mother has had diabetes at one or more of the registered pregnancies, and finally whether artificial insemination was required.

We also wish to control for effects related to wealth status, and thus include yearly after-tax *income* in 1,000 DKK, yearly *unemployment benefits* in 1,000 DKK and finally *home size* measured in square meters. Here, the values are those registered for the year of pregnancy. This category may be related to e.g. the ability to ensure good surroundings and a proper diet etc.

It has also previously been found that there is a linkage between birthweight outcomes and socio-economic factors such as marital status and level of education, see e.g. Abrevaya and Dahl (2008). We therefore include variables that indicate if the mother was *married* during the pregnancy period, the mother's *education* (a categorical variable summarised in Table 4, and included as dummy variables), and whether the mother was a *student* when pregnant. The latter variable may indicate how freely time can be organised and may proxy for how stressful workdays are.

The final two categories concern characteristics of the mother and child respectively. They include *height, weight*, and *age* of the mother (the latter two are also included in squares), and dummy variables for birth parity and the sex of the child.

Category	Description
0	No education.
1	Primary school (9 years compulsory, 1 year optional).
2	Secondary pre-university high school (3 years), or technical college, craftsmen, etc.(2–5 years).
3	College: short-cycle higher education programme (1–2 years).
4	College: medium-cycle higher education programme (3–4 years).
5	3-year academic (Bachelor) degree.
6	5-year academic (Master) degree.
7	PhD degree and above.

Table 4: Description of *education* categories.

A natural concern is whether or not there are relevant seasonality effects which should be controlled for. Buckles and Hungerman (2008) discuss whether the time of year affects birthweight and conclude that this is the case. They attribute such effects to a strong correlation with socio-economic characteristics, which are well represented in our data. Dehejia and Lleras-Muney (2004) investigate effects of unemployment rates on babies' health, and suggest that high unemployment is positively correlated with healthy babies. This is just one example of general year-specific phenomena which may have effects which are desirable to control for. In our analyses we do this by including birth-year dummy variables.

The overall choice of variables is in part motivated by previous studies such as Abrevaya and Dahl (2008) and Koenker and Hallock (2001). Some results are therefore comparable, and may confirm previous findings. To analyse the effect of maternal smoking, we use data on smoking both during and before pregnancy, allowing for smoke to have a causal effect in different ways. This approach differs from previous studies and relates to the discussion of whether "last-minute" intervention could be effective.

	1st child		2nd child		3rd child		4th child	
Variable	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
Birthweight	3503.82	(534.17)	3650.23	(531.20)	3674.22	(554.12)	3624.56	(547.75)
Smoked during	0.14		0.13		0.16		0.21	
Smoked before	0.30		0.24		0.25		0.28	
Drink	0.04		0.03		0.04		0.03	
Birth control pills	0.26		0.16		0.14		0.11	
Complications	0.22		0.25		0.28		0.26	
Doctor visits	3.07		3.00		2.94		2.80	
Prenatal visits	5.17		4.74		4.56		4.47	
Test tube baby	0.02		0.01		0.00		0.01	
Diabetes	0.01		0.01		0.02		0.03	
Income	107.04	(40.65)	135.62	(105.72)	149.39	(67.37)	151.94	(50.48)
Unemployment benefits	6.43	(17.41)	6.97	(18.57)	6.29	(17.98)	4.37	(14.52)
Home size	93.64	(45.20)	112.74	(44.26)	125.59	(43.71)	131.47	(44.26)
Married	0.43		0.65		0.76		0.72	
Student	0.23		0.14		0.09		0.08	
Education cat. 0	0.01		0.01		0.01		0.01	
Education cat. 1	0.13		0.12		0.17		0.29	
Education cat. 2	0.49		0.42		0.37		0.32	
Education cat. 3	0.05		0.06		0.04		0.03	
Education cat. 4	0.18		0.23		0.25		0.21	
Education cat. 5	0.05		0.04		0.03		0.02	
Education cat. 6	0.09		0.12		0.12		0.10	
Education cat. 7	0.00		0.01		0.01		0.01	
Height	168.56	(6.04)	168.57	(6.07)	168.20	(6.03)	167.36	(5.95)
Weight	63.92	(10.83)	65.24	(11.89)	65.79	(12.37)	65.71	(12.40)
Age	27.57	(3.75)	30.34	(3.82)	32.47	(3.83)	33.99	(4.18)
Male child	0.51		0.50		0.51		0.52	

	Birthweight quantiles			
Quantile	1st child	2nd child	3rd child	4th child
10%	2880	3030	3030	3002
25%	3200	3330	3350	3300
50%	3500	3650	3660	3650
75%	3850	4000	4020	3990
90%	4150	4300	4350	4288
Observations	6642	7416	2181	363

Table 5: Descriptive statistics for the Aarhus Birth Cohort.

4 Results

We consider our main estimation results to be the CRE estimates; especially those for the unbalanced panel and the CREM specification. From an economic policy perspective, the interpretation of the CRE estimates seems most appropriate: to the politician unobserved individual effects should be part of the unexplained ranking or distribution mechanism, yet for estimation purposes some notion of random assignment is called for to take into account the dependence between the unobserved and included covariates. Also, the CRE estimators are not cursed by incidental parameters. Their specifications give no reason why they should be, and indeed our simulations confirmed their good performance, also for short panels.

The FE estimators, on the other hand, are questionable in such a setting. We give some insight into some bounds for some available calibration options and argue that, even though the estimators are cursed, they seem to lead to conclusions that are difficult to refute.

Investigating the alleged negative effects of smoking behavior on birthweight outcomes is a prime objective in this analysis. Therefore we initiate the presentation of our results with a detailed discussion of the matter, using evidence from our battery of estimators, after which we elaborate on some of our other findings.

Table 6 summarizes the results for the *smoked during* variable as estimated by a variety of methods. These surely have one thing in common: a statement that prenatal smoking do not have negative direct effects on birthweight would be hard to justify, given the evidence from any of the estimators.

Consider first the results for the unbalanced data set.[7] Overall, the cross-section estimates provide the largest estimates (in absolute value) of the smoking effect. This holds, not only for the methods and calibrations shown here, but for all the many variations we have tried. The CREM "correction" estimates, S(CREM), absorbs an increasing part of the large effect alleged by a pure cross-section estimator, the further we move to the right in the birthweight distribution. This indicates that in the left tail, where we then have the largest adverse effect smoke during pregnancy, dependence between smoke and unobserved individual effects does not interfere much. Supposedly, there is more such "joint dependence" with birthweight at the larger quantiles. This supports a conclusion that smoking is more severe where it hurts the most: where babies are already prone to be low achievers when it comes to birthweight. In fact, all estimators except for 2SFE, lead to the conclusion that the adverse effect *relative* to birthweight is increasing to the left in the distribution when comparing point estimates to birthweight quantiles reported in Table 5. The 2SFE estimator here predicts a constant relative effect.

The KFE estimator can generally be calibrated to give results that lie between those of a pure dummy-variable estimator and a pure cross-section estimator. We will shortly discuss this in a little more detail. In our application, a pure dummy-variable regression is numerically infeasible. Letting $\lambda \to 0$ to approach the dummy-variable estimates, we learn that the effect of smoking decreases in absolute value. No value

[7] For all our estimations we have used a blocked pairwise subsampling bootstrap, as deemed appropriate by Abrevaya and Dahl (2008). The idea is that when sampling a mother, all her births are included to deal with the dependence in the observations.

Estimates for *smoked during*		Quantile Regressions				
		10%	25%	50%	75%	90%
Unbalanced data	CS	-188.063 ***	-181.629 ***	-169.165 ***	-177.479 ***	-200.441 ***
		(28.681)	(22.116)	(18.526)	(21.528)	(27.491)
	CREM	-190.485 ***	-112.107 ***	-75.991 ***	-90.337 ***	-2.081
		(49.047)	(35.800)	(27.952)	(34.573)	(49.101)
	S(CREM)	1.511	-81.205 *	-118.446 ***	-118.538 ***	-224.515 ***
		(60.811)	(47.264)	(36.936)	(44.050)	(57.950)
	KFE(5), $\lambda = 0.8$	-161.989 ***	-163.087 ***	-155.812 ***	-148.444 ***	-167.575 ***
		(26.249)	(18.478)	(16.592)	(19.996)	(24.541)
	KFE(5), $\lambda = 0.8$, post-est.	-162.821 ***	-156.776 ***	-146.783 ***	-186.265 ***	-183.513 ***
		(32.061)	(24.873)	(24.113)	(25.319)	(32.148)
	2SFE	-70.869 ***	-83.486 ***	-99.452 ***	-104.189 ***	-108.669 ***
		(26.305)	(21.007)	(19.840)	(20.379)	(24.320)
Balanced data	CREM	-247.002 ***	-154.444 ***	-88.071 ***	-110.895 ***	3.708
		(54.908)	(39.487)	(32.845)	(41.679)	(59.003)
	S(CREM)	76.550	-24.508	-98.878 **	-81.969	-214.324 ***
		(67.040)	(48.564)	(39.680)	(50.903)	(69.351)
	AD	-231.362 ***	-174.003 ***	-57.873	-126.478 ***	12.694
		(58.602)	(39.997)	(35.973)	(44.777)	(61.797)
	S_1(AD)	34.427	17.522	-52.118 *	19.774	-32.024
		(48.437)	(33.175)	(30.156)	(36.824)	(51.533)
	S_2(AD)	37.590	-21.545	-77.352 **	-88.525 **	-184.486 ***
		(48.435)	(37.411)	(34.246)	(41.459)	(54.086)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 6: Results for *smoked during* from a selection of estimators. The $S(\cdot)$-results are for the added CRE variables constructed from repeated measurements of *smoke during* (see the last part of Section 2.3.1). For the KFE estimates, λ refers to the penalty parameter, and "post. est" is where the model is re-estimated without penalty and zero-FE-terms.

of λ, however, leads to effects of smoking as small as claimed by the 2SFE estimator. The value of λ is non-negligible, and we are not aware of a practical rule for choosing it appropriately. However, our simulations confirm that the performance of the KFE estimator can be calibrated to outperform the 2SFE estimator. Since the latter acts like a lower bound in our case, there is strong evidence that there are significant direct adverse effects from smoke during pregnancy, also from a fixed effects perspective. However, there is much uncertainty about how adverse. The choice $\lambda = 0.8$, for which the results are reported in the table, penalizes to a degree where 30% of the mothers share intercept, and the remaining 70% are sufficiently different to get their own. Re-estimating the selected model has a more pronounced effect in the right tail. An argument for re-estimation is that the penalty affects the non-zero FE terms, whereas an argument against re-estimation is to preserve some degree of control over the size of estimated individual effects.

In the lower part of Table 6, we provide CRE results from the smaller balanced panel. We include this mainly to show that the AD and CREM specifications give very similar estimates, justifying the more simple CREM specification of S which then allows for the inclusion of more observations given by the unbalanced panel. The point estimates for the balanced panel indicate a slightly larger effect, compared to the unbalanced one, even though we cannot deem them statistically different. One explanation, however, could be that mothers who give birth to unhealthy babies (in terms of low birthweight) choose not to get a third or fourth child, this resulting in some kind of sample selection issue. This would again point to the unbalanced panel as the more appropriate.

A general point that we need to emphasize is that even at the first decile births are not categorised as low birthweight (often defined as 2,500 grams), cf. Table 5. Unfortunately, it is not possible to obtain reasonable results for lower quantiles due to the very few extreme observations. However, it does not seem reasonable to expect the adverse effects of smoking to diminish as we move into the extreme left of the distribution. In fact, the trend in the CRE models suggests exactly the opposite. To illustrate the trend visually we have plotted point CRE estimates for a whole range of quantiles in Figure 3 (left panel). For comparison we show some FE estimates in the right panel. As KFE(k) estimations are problematic for such a "grid", we use a KFE(1) specification. The right panel, where we also include the cross-section estimates, also serves to illustrate how estimates move as a function of λ.

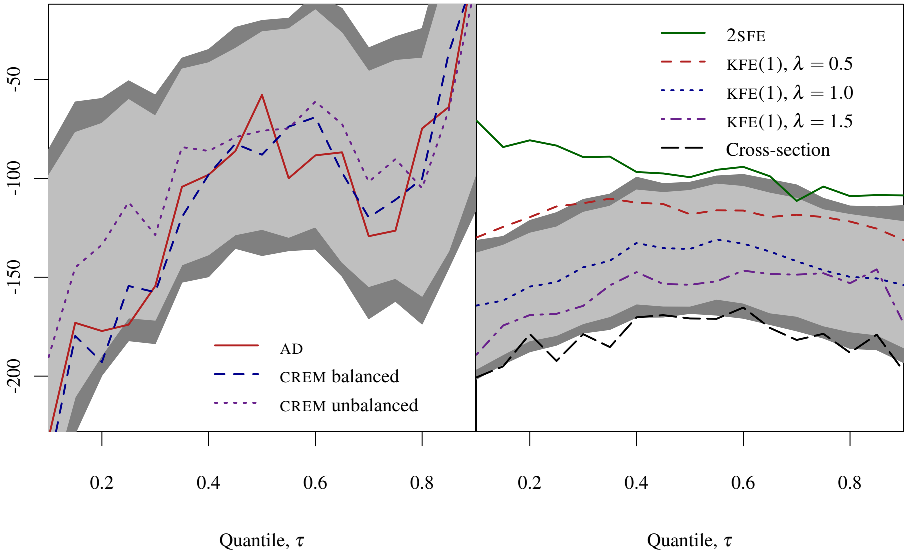

Fig. 3: *Smoke during* estimates plotted for a grid of quantile indices and estimators. The gray areas show the bootstrapped 90% and 95% confidence intervals.

To add a little insight into what happens when varying λ, we consider the KFE(5) point estimates for *smoked during* and the ratio of FE terms shrunken to zero as a function of λ. We present this sensitivity analysis in Figure 4. The top panel shows how the effect decreases at all quantiles as we penalize more. The trend stops at $\lambda \approx 1.6$ where the bottom panel shows that almost all FE terms are set to zero. We also show the penalized ratio for KFE(1). It seems that FE terms are (fully) affected by the penalty in chunks. This feature is most pronounced for KFE(1), and it appears that this effect is smoothed out for KFE(k) as k increases.

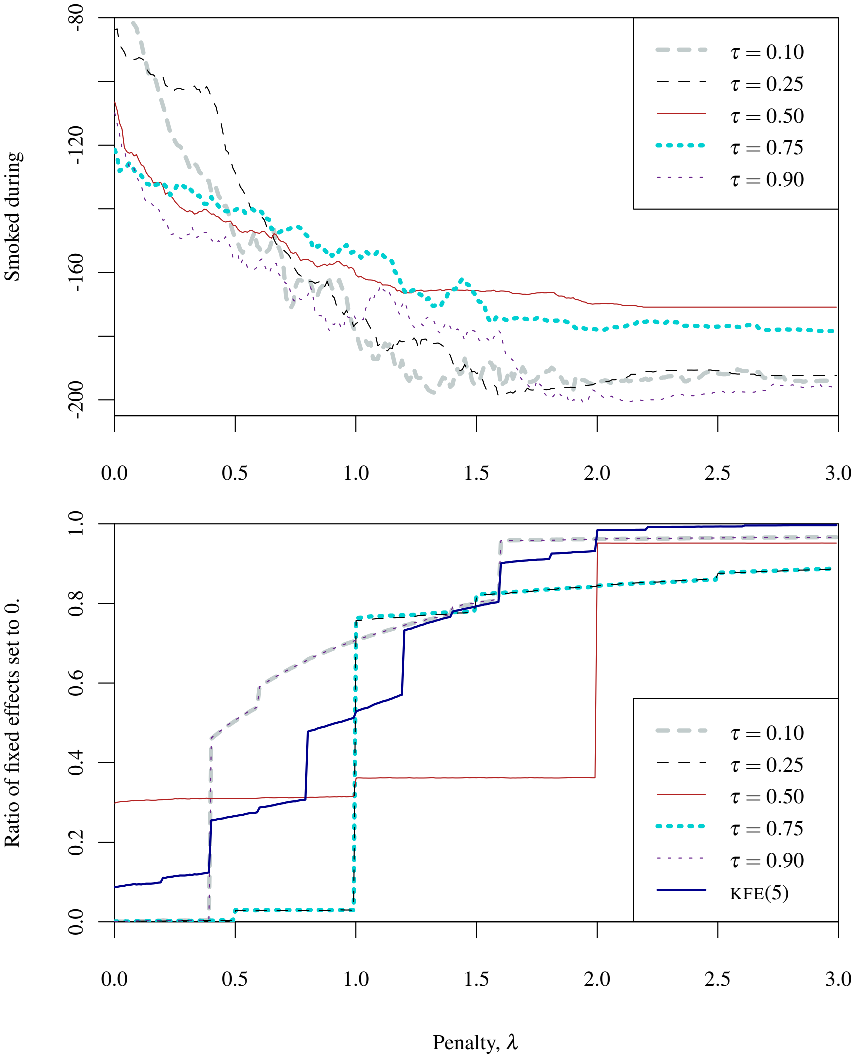

Fig. 4: The top panel shows the KFE(5) estimates for *smoked during* as the penalty parameter λ varies. The bottom panel shows the ratio of FE-terms shrunken to zero because of the penalty as a function of λ.

Our analysis, of course, also includes data on whether the mothers *smoked before* their pregnancies. Interestingly, however, this seems to be unimportant when it comes to birthweight: no estimators find noteworthy significance. We therefore do not present a similar discussion for this variable, although some results can be found in Appendix B. We find the absence of statistical significance a very important and quite comforting result. It suggests that smoking behavior prior to pregnancy is not crucial for birthweight outcome, and it gives future mothers who are smoking a very good argument to quit in time. Intervention policy can therefore prove profitable, as reducing the number of low birthweight babies reduces both monetary as well as socio-economic costs.

Maternal smoking is also one of the prime interests in the study by Abrevaya and Dahl (2008), which is based on American data, more specifically natality data from Washington and Arizona. In their study they only have a smoke variable comparable to our *smoked during*, but for this variable they also find significant negative effects. However, their findings suggest somewhat less severe effects, in the range around -80 to -60 grams. Whether this can in fact be attributed to actual differences or is a consequence of measurement error is hard to say, but the latter could perhaps be attributed to smoking being less of a taboo in Denmark, which could lead to a lower degree of misreporting. Furthermore, they find that cross sectional estimates exaggerate the effect of smoking even more than in this case.

In addition to the variables pertaining to maternal smoking our analysis also contains a number of other interesting variables. A set of results for the CREM model using the unbalanced dataset is given in Table 7. Again, we will focus on these results and only consider those from the other models on a few occasions. A complete set of results for all the models is available as a separate appendix, and the most relevant results are listed in Appendix B. For the CRE-specifications we use all variables that vary from birth to birth to construct S_m, except education and year dummy-variables. The variables *height* and *diabetes* do not vary in our sample and are therefore not used, either for S_m or in the fixed effects model.

One behavioural aspect which receives great attention, especially related to pregnancies, and which is the subject of much societal debate, is drinking habits. This is also a topic where policy is conducted in the attempt to affect people's behaviour, e.g. taxes and age restrictions on the purchase of alcohol. In the light of the strong belief that alcohol intake during pregnancy has negative health implications on the foetus, it seems puzzling that these estimations show no significance in the CREM model and no or very little in the other models considered. There could be many reasons why no significance shows up in our results, one of which could be measurement or reporting errors. Even so, drinking may be the cause of many other health implications which are not related to birthweight.

In the behaviour category we also have the variable *birth control pills*. As mentioned earlier this can be thought to proxy for whether the pregnancy was planned or not. We do see some signs of significance for this variable, but the extent of this varies between models. However, in general the point estimates are negative as would be expected if it indeed acts as a proxy for unwanted pregnancies.

In the health category of variables we have three variables which show clearly significant effects. *Prenatal visits* are of special interest since it is a preventive measure

	Quantile Regressions					
	10%	25%	50%	75%	90%	OLS
Smoked during	-190.485 ***	-112.107 ***	-75.991 ***	-90.337 ***	-2.081	-94.897 ***
	(49.047)	(35.800)	(27.952)	(34.573)	(49.101)	(23.618)
Birth control pills	-27.853	-52.478 ***	-28.928 **	-21.657	-19.116	-33.813 ***
	(25.097)	(18.130)	(14.118)	(16.678)	(22.708)	(11.650)
Complications	-122.386 ***	-67.361 ***	-46.073 ***	-29.562 **	-11.029	-65.723 ***
	(24.632)	(16.595)	(12.271)	(14.727)	(21.520)	(10.636)
Prenatal visits	109.582 ***	90.162 ***	77.477 ***	80.245 ***	73.484 ***	64.011 ***
	(11.583)	(5.906)	(4.583)	(5.080)	(6.969)	(11.670)
Test tube baby	-2.683	64.214	30.674	-57.277	-251.771 ***	-34.417
	(108.153)	(70.387)	(61.737)	(63.434)	(86.622)	(46.240)
Diabetes	181.554 *	217.704 ***	280.896 ***	315.398 ***	373.631 ***	282.529 ***
	(94.004)	(70.214)	(55.456)	(62.333)	(88.341)	(54.962)
Student	20.743	4.311	1.137	10.260	51.933 *	17.787
	(31.961)	(22.674)	(21.060)	(22.319)	(29.430)	(15.434)
Height	6.662 ***	8.423 ***	10.489 ***	11.419 ***	11.183 ***	10.269 ***
	(1.525)	(1.088)	(0.975)	(1.074)	(1.320)	(0.906)
Weight	5.722	2.412	-8.622	2.040	23.161 **	7.950
	(14.546)	(9.823)	(8.403)	(8.182)	(11.050)	(6.514)
Weight^2	-0.026	0.008	0.063	-0.010	-0.127 *	-0.036
	(0.099)	(0.066)	(0.058)	(0.054)	(0.071)	(0.043)
Age	-40.769	-24.258	-19.409	-29.653 *	-0.204	-14.971
	(30.152)	(20.280)	(16.161)	(17.768)	(26.393)	(14.840)
Age^2	0.867 *	0.516	0.402	0.465	0.065	0.333
	(0.467)	(0.319)	(0.256)	(0.285)	(0.420)	(0.236)
Second child	177.725 ***	162.143 ***	150.112 ***	175.701 ***	156.739 ***	162.162 ***
	(21.562)	(14.632)	(12.540)	(14.267)	(19.457)	(11.515)
Third child	213.334 ***	191.121 ***	190.076 ***	252.706 ***	231.066 ***	209.153 ***
	(33.522)	(24.012)	(22.312)	(24.301)	(31.575)	(20.485)
Fourth child	172.238 ***	182.380 ***	181.915 ***	233.246 ***	202.214 ***	191.631 ***
	(56.800)	(41.156)	(37.683)	(40.316)	(51.240)	(34.217)
Male child	121.921 ***	128.415 ***	137.299 ***	162.926 ***	186.615 ***	145.904 ***
	(16.220)	(12.039)	(10.115)	(12.611)	(16.705)	(8.042)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 7: Results for the CREM estimation using the unbalanced data set. Insignificant variables are not reported here but in Appendix B. These are: *smoked before, drink, doctor visits, income, unemployment benefits, home size, married,* and all education categories. Dummy variables for birth year are mostly significant but removed from this table in the interest of space. Results for the constructed CRE variables (i.e. those in S_m) can also be found in Appendix B. The OLS estimates are from a Mundlak regression with the projection being the same as S_m.

intended to ensure good health of the foetus, and therefore its effect is of great interest to policy makers. The main problem with *prenatal visits*, however, is that there may be two reasons for consulting a midwife, either as a routine/precautionary measure or because of complications. It is not possible to directly distinguish between these two effects of the variable. Therefore the estimations also include *complications*, which in part controls for this, thus leaving us with the preventive effect. We see that *prenatal visits* are significant, indicating a positive preventive effect. Further, *complications* have a significant negative effect as would be expected. However, this variable is problematic, as indicated in the last section, since it includes cases of intrauterine growth restrictions, which may be one of the channels through which smoking reduces birthweight. It is not possible to separate out this part of the variable, and consequently as a robustness check regressions have been run without *complications*. The resulting outcome had only minor changes in the point estimates and did not alter any conclusions. On this basis it is concluded that the prevalence of intrauterine growth restrictions does not constitute a problem for the interpretation of the results, in particular those for *smoked during*.

The last significant variable in this category is *diabetes*, which has a positive effect on the right tail of the birthweight distribution. This is in accordance with the medical literature, where diabetes is commonly accepted as a birthweight-increasing factor. Finally, both *doctor visits* and *test tube baby* show only little significance in the CREM model. In a few cases, they show moderate significance in the fixed effects models. The former can be thought of as a general measure of the mother's health. However, it is hard to say how good a proxy it really is, since it represents, not only birth-related health, but also general illness or even hypochondria. The fact that the latter is mostly insignificant need not say anything about causality, but may be due to a very small number of test tube babies in the sample.

The variables in the wealth and socio-economic categories are in general all insignificant. We do, however, see two exceptions. First, the variable *student* is mostly significant in the fixed effects models, which is in contrast to the CRE models. Second, the education variables do show moderate signs of significance in some of the fixed effects specifications, but there is not any general consensus on the significance between the models. That these categories are largely unimportant is not particularly unexpected when considering the welfare system in Denmark, where the social benefits available in general (and to mothers in particular) are quite generous. This is in contrast to the results from e.g. Abrevaya and Dahl (2008). They find, for instance, that marital status is highly significant. A reason for this difference could be that America has a substantial social gap compared to Denmark. This will undoubtedly have consequences for unmarried mothers in America, who do not have the same social benefits as offered in Denmark. Another, perhaps more subtle reason could be the extent to which marriage can proxy for unobserved characteristics or ability of women. The choice of why and when to get married may be culturally dependent, which is supported by the descriptive statistics. There is quite a difference in proportions of pregnancies in and out of wedlock in their American data and our Danish data, which suggests that it is more uncommon to have children out of wedlock in America. When combined, these arguments may be used to explain why the American data suggest that marriage has a positive effect and no such evidence is found in the Danish data.

Abrevaya and Dahl also find that education has a significant effect, while we find little evidence of such an effect. This could very well be due to the costs associated with education in America. This is in contrast to Denmark where education is free. The variable may therefore proxy for wealth status which, as argued before, seems irrelevant in Denmark.

The mother's characteristics are largely insignificant in the "main" terms of the CREM specification (those in X_{mb}). The augmented CRE terms, however, do show some significance (those in S_m). This could be interpreted as a "part" of the heterogeneity, e.g. for *weight*: the overall stature of the mother can be more important than birth-specific fluctuations. *Height*, is highly significant. However, no extra CRE variable is constructed for height because it is birth-invariant. The fact that the *height* and *weight* variables are significant, for one part of the specification or the other, seems very natural. What is puzzling, though, is that *age* does not appear to have much significance. Abrevaya and Dahl (2008) find a significant effect of *age*, and the literature suggests that there is an optimal age, see e.g. Royer (2004).

For these variables the fixed effects models differ considerably. First, because *height* is birth invariant it cannot be included in these models and should instead be captured by the fixed effect. Second, *weight* is in most cases significant. This is to be expected as it cannot be captured by the fixed effects, but will most likely affect birthweight. Finally, *age* does show moderate signs of significance in some of the fixed effects specifications, but not to a degree where we are confident enough to draw any firm conclusions.

Regarding child characteristics, we find that the parity variables *second, third* and *fourth child,* and *male child* are significant and positive across all quantiles. This is to be expected since it is generally acknowledged that the birthweight of male children is higher on average, and that birthweight increases with parity of the mother. This confirms the results of previous studies.

5 Concluding remarks

In this paper we have found strong evidence that smoking during pregnancy has adverse consequences for birthweight outcomes. The documented connection between babies' birthweight and their overall health, along with the costs associated with low birthweight, makes this a very important result. The effect appears to worsen the further one moves to the left in the birthweight distribution, especially when measured relative to birthweight at the corresponding quantiles.

The significant effect of smoking has been documented before, but we add to these results in several ways. The richness of the applied data set allowed us to control for many potentially important characteristics which were not included in previous studies. Furthermore, we use several estimators and provide a detailed discussion of their differences in interpretation and performance. Given the results from this battery of estimators, the adverse effect of smoking on birthweight seems irrefutable, regardless of estimation approach and which of the two discussed interpretations is desired for the estimated coefficients.

As icing on the cake, our analysis used information on smoking behavior prior to pregnancy, allowing for a separation of effects. Only smoking during pregnancy has a pronounced significant effect, a result speaking for intervention campaigns as a worthwhile activity.

Acknowledgements We are grateful Jason Abrevaya, Niels Haldrup, Marianne Simonsen, Chiara Orsini, Ulrich Kaiser, Tine B. Henriksen, and Kirsten Wisborg for helpful comments. Lise V. Hansen offered invaluable assistance by providing access to the data from Aarhus University Hospital, Skejby. The terms of the data-sharing agreements do not allow release of these data. However, codes for estimation in R are available from the corresponding author upon request. We would also also like to thank Roger Koenker for useful discussion and for writing the quantreg package for R that was used as starting point for the quantile estimations in this paper. The authors gratefully acknowledge financial support from the Danish Social Sciences Research Council (grant 275-06-0105) and the Center for Research in Econometric Analysis of Time Series, CREATES, funded by The Danish National Research Foundation.

A Simulation results

		$B=2$		$B=3$		$B=5$	
	τ	$M=499$	$M=999$	$M=499$	$M=999$	$M=499$	$M=999$
Cross-section FE target	0.25	−0.3380	−0.3402	−0.3412	−0.3443	−0.3435	−0.3405
		(0.3713)	(0.3561)	(0.3618)	(0.3538)	(0.3560)	(0.3467)
	0.50	−0.3724	−0.3740	−0.3755	−0.3718	−0.3718	−0.3722
		(0.3968)	(0.3867)	(0.3920)	(0.3808)	(0.3813)	(0.3770)
	0.75	−0.4321	−0.4344	−0.4310	−0.4282	−0.4186	−0.4245
		(0.4598)	(0.4468)	(0.4489)	(0.4373)	(0.4301)	(0.4303)
Dummy regression	0.25	0.6717	0.6626	0.3531	0.3474	0.1438	0.1431
		(0.6980)	(0.6770)	(0.3813)	(0.3619)	(0.1763)	(0.1595)
	0.50	−0.0028	−0.0119	−0.0047	−0.0055	−0.0035	−0.0053
		(0.1898)	(0.1395)	(0.1272)	(0.0904)	(0.0865)	(0.0645)
	0.75	−0.6773	−0.6864	−0.3596	−0.3598	−0.1438	−0.1475
		(0.7033)	(0.7003)	(0.3877)	(0.3736)	(0.1752)	(0.1645)
KFE3	0.25	−0.0873	−0.0958	−0.0493	−0.0513	−0.0927	−0.0911
		(0.1872)	(0.1486)	(0.1283)	(0.0981)	(0.1322)	(0.1129)
	0.50	−0.1391	−0.1425	−0.1534	−0.1530	−0.1426	−0.1444
		(0.2089)	(0.1818)	(0.1887)	(0.1722)	(0.1641)	(0.1560)
	0.75	−0.3664	−0.3694	−0.2452	−0.2396	−0.1828	−0.1889
		(0.3977)	(0.3852)	(0.2753)	(0.2547)	(0.2062)	(0.2009)
KFE3 Post est.	0.25	0.0877	0.0807	0.1105	0.1081	0.0614	0.0600
		(0.1967)	(0.1425)	(0.1682)	(0.1421)	(0.1157)	(0.0905)
	0.50	−0.0224	−0.0321	−0.0474	−0.0430	−0.0253	−0.0269
		(0.1821)	(0.1357)	(0.1264)	(0.0951)	(0.0875)	(0.0676)
	0.75	−0.2531	−0.2582	−0.2052	−0.2002	−0.1040	−0.1086
		(0.3030)	(0.2844)	(0.2476)	(0.2207)	(0.1431)	(0.1279)
KFE1	0.25	−0.1993	−0.2071	−0.1275	−0.1322	−0.1027	−0.1006
		(0.2582)	(0.2368)	(0.1725)	(0.1542)	(0.1395)	(0.1207)
	0.50	−0.2239	−0.2315	−0.1827	−0.1818	−0.1331	−0.1344
		(0.2615)	(0.2520)	(0.2144)	(0.1989)	(0.1570)	(0.1477)
	0.75	−0.3690	−0.3749	−0.2996	−0.2943	−0.2865	−0.2934
		(0.3980)	(0.3902)	(0.3247)	(0.3066)	(0.3009)	(0.3008)
2SFE	0.25	0.3921	0.3844	0.2598	0.2551	0.1563	0.1553
		(0.4285)	(0.4014)	(0.2882)	(0.2707)	(0.1844)	(0.1687)
	0.50	−0.0020	−0.0110	−0.0208	−0.0210	−0.0259	−0.0267
		(0.1640)	(0.1169)	(0.1162)	(0.0844)	(0.0879)	(0.0664)
	0.75	−0.4036	−0.4080	−0.3004	−0.2981	−0.1961	−0.2005
		(0.4355)	(0.4249)	(0.3262)	(0.3101)	(0.2155)	(0.2104)
Cross-section CRE target	0.25	−0.3825	−0.3848	−0.3858	−0.3889	−0.3881	−0.3851
		(0.4123)	(0.3989)	(0.4042)	(0.3973)	(0.3992)	(0.3906)
	0.50	−0.3724	−0.3740	−0.3755	−0.3718	−0.3718	−0.3722
		(0.3968)	(0.3867)	(0.3920)	(0.3808)	(0.3813)	(0.3770)
	0.75	−0.3875	−0.3898	−0.3864	−0.3836	−0.3740	−0.3799
		(0.4182)	(0.4036)	(0.4063)	(0.3937)	(0.3869)	(0.3864)
CREM	0.25	−0.0164	−0.0253	−0.0225	−0.0247	−0.0210	−0.0185
		(0.2063)	(0.1416)	(0.1439)	(0.1029)	(0.1020)	(0.0739)
	0.50	−0.0004	−0.0092	−0.0103	−0.0073	−0.0062	−0.0071
		(0.1749)	(0.1237)	(0.1290)	(0.0916)	(0.0904)	(0.0638)
	0.75	−0.0018	−0.0052	0.0048	0.0057	0.0144	0.0093
		(0.2017)	(0.1398)	(0.1431)	(0.1030)	(0.1045)	(0.0745)
AD	0.25	−0.0126	−0.0247	−0.0217	−0.0227	−0.0191	−0.0181
		(0.2028)	(0.1399)	(0.1432)	(0.1029)	(0.1021)	(0.0728)
	0.50	0.0005	−0.0093	−0.0086	−0.0077	−0.0066	−0.0071
		(0.1723)	(0.1237)	(0.1287)	(0.0919)	(0.0905)	(0.0635)
	0.75	−0.0066	−0.0060	0.0028	0.0062	0.0127	0.0095
		(0.2018)	(0.1391)	(0.1424)	(0.1026)	(0.1046)	(0.0749)

Table 8: Bias and root mean squared error (rmse) for simulation of (23) with $\gamma = 0$, i.e. no scale effect of the individual effects.

	τ	$B=2$ $M=499$	$B=2$ $M=999$	$B=3$ $M=499$	$B=3$ $M=999$	$B=5$ $M=499$	$B=5$ $M=999$
Cross-section FE target	0.25	−0.1453	−0.1508	−0.1465	−0.1493	−0.1482	−0.1487
		(0.1853)	(0.1688)	(0.1739)	(0.1641)	(0.1671)	(0.1591)
	0.50	−0.4118	−0.4218	−0.4112	−0.4131	−0.4159	−0.4170
		(0.4326)	(0.4307)	(0.4252)	(0.4205)	(0.4240)	(0.4213)
	0.75	−0.5625	−0.5662	−0.5608	−0.5582	−0.5664	−0.5645
		(0.5855)	(0.5777)	(0.5774)	(0.5665)	(0.5753)	(0.5698)
Dummy regression	0.25	0.6746	0.6619	0.4037	0.4078	0.2317	0.2344
		(0.6973)	(0.6714)	(0.4241)	(0.4179)	(0.2482)	(0.2430)
	0.50	−0.0000	−0.0127	−0.0027	0.0003	−0.0020	−0.0005
		(0.1766)	(0.1134)	(0.1126)	(0.0826)	(0.0761)	(0.0573)
	0.75	−0.6744	−0.6871	−0.4076	−0.4071	−0.2327	−0.2313
		(0.6971)	(0.6963)	(0.4270)	(0.4179)	(0.2478)	(0.2395)
KFE3	0.25	0.1404	0.1339	0.2401	0.2387	0.1974	0.2002
		(0.1979)	(0.1605)	(0.2655)	(0.2514)	(0.2160)	(0.2099)
	0.50	−0.1861	−0.1970	−0.1769	−0.1756	−0.1575	−0.1571
		(0.2367)	(0.2179)	(0.2049)	(0.1910)	(0.1745)	(0.1667)
	0.75	−0.5255	−0.5303	−0.4710	−0.4708	−0.4405	−0.4393
		(0.5514)	(0.5426)	(0.4889)	(0.4803)	(0.4517)	(0.4452)
KFE3 Post est.	0.25	0.3908	0.3829	0.3568	0.3556	0.3281	0.3289
		(0.4276)	(0.3999)	(0.3795)	(0.3667)	(0.3424)	(0.3363)
	0.50	−0.0235	−0.0364	−0.0378	−0.0369	−0.0209	−0.0185
		(0.1717)	(0.1140)	(0.1137)	(0.0866)	(0.0778)	(0.0593)
	0.75	−0.5240	−0.5309	−0.4238	−0.4220	−0.3684	−0.3682
		(0.5515)	(0.5447)	(0.4442)	(0.4331)	(0.3797)	(0.3745)
KFE1	0.25	−0.0206	−0.0255	0.0462	0.0446	0.0653	0.0668
		(0.1284)	(0.0848)	(0.1107)	(0.0834)	(0.1002)	(0.0874)
	0.50	−0.2401	−0.2503	−0.1690	−0.1690	−0.1137	−0.1118
		(0.2737)	(0.2633)	(0.1992)	(0.1855)	(0.1363)	(0.1257)
	0.75	−0.5009	−0.5108	−0.4130	−0.4143	−0.4239	−0.4217
		(0.5255)	(0.5229)	(0.4299)	(0.4235)	(0.4341)	(0.4272)
2SFE	0.25	0.5075	0.5054	0.3986	0.3962	0.2813	0.2839
		(0.5331)	(0.5167)	(0.4157)	(0.4048)	(0.2940)	(0.2902)
	0.50	−0.0004	−0.0091	−0.0226	−0.0220	−0.0273	−0.0270
		(0.1585)	(0.1052)	(0.1112)	(0.0823)	(0.0799)	(0.0611)
	0.75	−0.5913	−0.5983	−0.5150	−0.5165	−0.4201	−0.4197
		(0.6162)	(0.6104)	(0.5314)	(0.5256)	(0.4294)	(0.4250)
Cross-section CRE target	0.25	−0.2084	−0.2139	−0.2095	−0.2124	−0.2113	−0.2118
		(0.2380)	(0.2269)	(0.2296)	(0.2230)	(0.2249)	(0.2192)
	0.50	−0.3865	−0.3965	−0.3859	−0.3878	−0.3906	−0.3918
		(0.4087)	(0.4060)	(0.4008)	(0.3957)	(0.3992)	(0.3963)
	0.75	−0.5083	−0.5120	−0.5066	−0.5040	−0.5122	−0.5104
		(0.5336)	(0.5247)	(0.5250)	(0.5132)	(0.5220)	(0.5162)
CREM	0.25	−0.0061	−0.0116	−0.0107	−0.0126	−0.0256	−0.0243
		(0.1600)	(0.1084)	(0.1138)	(0.0802)	(0.0853)	(0.0649)
	0.50	−0.0217	−0.0310	−0.0118	−0.0095	−0.0066	−0.0081
		(0.1890)	(0.1255)	(0.1254)	(0.0945)	(0.0911)	(0.0657)
	0.75	−0.0002	−0.0084	−0.0031	−0.0048	−0.0144	−0.0103
		(0.2146)	(0.1518)	(0.1538)	(0.1107)	(0.1075)	(0.0784)
AD	0.25	−0.0076	−0.0137	−0.0129	−0.0132	−0.0232	−0.0231
		(0.1586)	(0.1056)	(0.1161)	(0.0803)	(0.0842)	(0.0646)
	0.50	−0.0201	−0.0271	−0.0115	−0.0094	−0.0059	−0.0089
		(0.1889)	(0.1243)	(0.1274)	(0.0939)	(0.0901)	(0.0659)
	0.75	0.0010	−0.0063	−0.0017	−0.0053	−0.0149	−0.0103
		(0.2167)	(0.1513)	(0.1491)	(0.1102)	(0.1072)	(0.0792)

Table 9: Bias and root mean squared error (rmse) for simulation of (23) with $\gamma = 1$, i.e. the individual effects have scale effects.

B Empirical results

	Quantile Regressions					
	10%	25%	50%	75%	90%	OLS
Smoked during	-231.362 ***	-174.003 ***	-57.873	-126.478 ***	12.694	-109.920 ***
	(58.602)	(39.997)	(35.973)	(44.777)	(61.797)	(27.174)
Smoked before	-14.954	-9.730	-14.592	-21.266	-32.916	-17.915
	(45.387)	(32.215)	(26.704)	(32.912)	(47.865)	(21.165)
Drink	2.451	-14.800	-34.876	-70.408 *	-59.946	-56.370 *
	(62.183)	(46.010)	(37.541)	(40.433)	(54.935)	(30.156)
Birth control pills	-34.450	-53.682 ***	-30.988 *	-19.569	-22.692	-33.379 ***
	(27.588)	(18.999)	(17.313)	(19.016)	(27.585)	(12.648)
Complications	-100.794 ***	-61.137 ***	-26.939 *	-23.624	2.355	-54.648 ***
	(30.312)	(20.413)	(14.485)	(17.466)	(24.221)	(13.054)
Doctor visits	6.119	18.112 *	16.192 **	11.760	8.599	15.857 *
	(12.870)	(9.804)	(8.155)	(8.255)	(11.528)	(8.399)
Prenatal visits	92.698 ***	84.246 ***	74.827 ***	74.228 ***	79.509 ***	55.543 ***
	(15.312)	(8.311)	(5.941)	(5.990)	(7.938)	(13.689)
Test tube baby	20.611	-27.565	-0.812	-40.657	-181.511 *	-54.088
	(127.855)	(79.473)	(65.856)	(73.092)	(97.172)	(50.084)
Diabetes	103.043	150.006 *	256.696 ***	285.550 ***	282.415 ***	221.760 ***
	(121.416)	(84.986)	(70.130)	(63.294)	(95.084)	(68.298)
Income	-0.080	-0.016	-0.122	-0.048	-0.100	-0.114
	(0.365)	(0.270)	(0.221)	(0.255)	(0.347)	(0.178)
Unemployment benefits	-0.430	-0.112	0.034	-0.220	-0.190	-0.328
	(0.612)	(0.457)	(0.373)	(0.428)	(0.596)	(0.298)
Home size	-0.255	-0.316	0.138	0.115	0.039	-0.114
	(0.275)	(0.219)	(0.204)	(0.255)	(0.306)	(0.154)
Married	-2.424	-2.771	-3.197	13.586	-29.501	9.930
	(32.435)	(24.068)	(18.997)	(22.819)	(30.852)	(15.214)
Student	1.707	5.730	9.500	27.206	38.908	11.019
	(38.078)	(25.659)	(21.703)	(25.454)	(35.496)	(16.583)
Height	6.879 ***	9.110 ***	10.965 ***	11.956 ***	11.858 ***	10.711 ***
	(1.676)	(1.242)	(1.095)	(1.185)	(1.496)	(1.009)
Weight	14.741	6.959	8.628	11.938	18.380	14.603 *
	(19.519)	(13.179)	(11.240)	(11.493)	(14.762)	(8.772)
$Weight^2$	-0.088	-0.016	-0.051	-0.080	-0.092	-0.085
	(0.134)	(0.091)	(0.077)	(0.078)	(0.099)	(0.060)
Age	-20.490	-14.207	-44.707 **	-55.060 **	-59.251	-51.252 ***
	(44.525)	(29.026)	(21.477)	(24.239)	(36.313)	(18.965)
Age^2	0.446	0.272	0.801 **	0.940 **	1.034 *	0.943 ***
	(0.689)	(0.463)	(0.348)	(0.390)	(0.588)	(0.303)
Second child	172.737 ***	176.777 ***	161.944 ***	167.637 ***	173.776 ***	161.824 ***
	(38.773)	(25.237)	(20.194)	(22.451)	(32.241)	(18.413)
Male child	122.307 ***	143.390 ***	142.821 ***	174.666 ***	204.769 ***	152.488 ***
	(19.951)	(14.244)	(12.124)	(13.834)	(18.670)	(9.888)
Education Cat. 1	89.874	-88.905	-50.289	14.492	47.121	-42.019
	(117.650)	(82.602)	(59.506)	(65.153)	(79.828)	(57.275)
Education Cat. 2	127.210	-45.128	10.293	77.617	140.297 *	30.576
	(112.123)	(80.818)	(58.588)	(63.556)	(78.042)	(56.087)
Education Cat. 3	143.775	-16.251	-9.094	68.587	129.169	30.857
	(119.275)	(82.667)	(62.246)	(68.082)	(84.486)	(59.790)
Education Cat. 4	142.865	-34.228	19.312	110.734 *	141.894 *	43.511
	(113.361)	(80.779)	(59.365)	(65.097)	(80.660)	(58.183)
Education Cat. 5	179.570	-1.531	14.288	91.657	168.319 *	54.905
	(119.177)	(81.216)	(61.821)	(67.131)	(86.391)	(60.299)
Education Cat. 6	168.811	-1.022	30.895	91.665	156.424 *	61.238
	(115.427)	(83.092)	(60.580)	(65.951)	(83.129)	(58.816)
Education Cat. 7	214.014	23.381	38.629	154.758	123.545	68.702
	(160.280)	(120.056)	(95.431)	(104.422)	(110.882)	(89.514)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 10: Estimation results from the AD model using the balanced dataset. Main variables.

	Quantile Regressions					
	10%	25%	50%	75%	90%	OLS
Smoked during (i)	34.427	17.522	-52.118 *	19.774	-32.024	-17.594
	(48.437)	(33.175)	(30.156)	(36.824)	(51.533)	(26.501)
Smoked during (ii)	37.590	-21.545	-77.352 **	-88.525 **	-184.486 ***	-71.036 **
	(48.435)	(37.411)	(34.246)	(41.459)	(54.086)	(31.282)
Smoked before (i)	21.684	28.449	12.392	-6.865	18.485	18.441
	(33.333)	(26.032)	(22.050)	(26.626)	(38.512)	(18.975)
Smoked before (ii)	-4.789	16.983	12.363	52.183 *	54.067	23.716
	(39.965)	(30.875)	(27.373)	(30.820)	(42.059)	(23.153)
Drink (i)	8.472	-15.287	-9.150	17.372	-7.084	14.925
	(45.042)	(43.505)	(37.213)	(42.717)	(52.395)	(30.150)
Drink (ii)	-43.062	17.824	54.152	12.678	-6.692	34.730
	(55.961)	(45.819)	(35.595)	(39.593)	(59.906)	(35.509)
Birth control pills (i)	-7.967	12.434	-12.738	-17.245	-26.740	-13.656
	(25.764)	(18.860)	(16.818)	(18.882)	(25.430)	(13.801)
Birth control pills (ii)	34.883	50.037 **	18.906	20.941	32.899	37.296 **
	(27.015)	(20.325)	(17.742)	(18.587)	(28.468)	(15.719)
Complications (i)	-65.146 **	-26.441	-22.255	-17.894	-11.197	-37.698 **
	(26.855)	(18.963)	(14.585)	(18.625)	(24.055)	(15.105)
Complications (ii)	-88.984 ***	-55.124 ***	-31.961 **	-31.331 *	-42.418 *	-57.531 ***
	(27.224)	(20.279)	(15.756)	(16.976)	(24.115)	(14.989)
Doctor visits (i)	4.390	-2.316	1.404	-1.955	6.506	9.751
	(12.360)	(9.992)	(8.851)	(9.915)	(13.147)	(8.532)
Doctor visits (ii)	1.283	-3.983	-1.185	-4.714	-8.715	-3.847
	(12.543)	(9.229)	(8.435)	(8.106)	(11.210)	(7.503)
Prenatal visits (i)	29.114 **	30.386 ***	36.961 ***	38.947 ***	30.897 ***	27.070 **
	(12.883)	(9.344)	(6.827)	(6.081)	(7.713)	(11.495)
Prenatal visits (ii)	0.370	-0.392	2.953	3.607	-0.300	0.848
	(9.028)	(5.291)	(4.562)	(5.091)	(5.969)	(6.655)
Test tube baby (i)	46.405	32.526	44.364	17.930	91.791	62.780
	(108.364)	(65.722)	(53.476)	(72.293)	(110.727)	(55.740)
Test tube baby (ii)	-41.714	20.727	-30.053	78.943	147.260 *	38.015
	(101.538)	(82.967)	(77.418)	(84.050)	(82.467)	(62.333)
Income (i)	0.211	0.154	-0.276	-0.390	-0.390	-0.143
	(0.328)	(0.268)	(0.243)	(0.282)	(0.365)	(0.213)
Income (ii)	0.110	0.005	0.082	0.047	-0.001	0.083
	(0.326)	(0.245)	(0.194)	(0.198)	(0.266)	(0.167)
Unemployment benefits (i)	-0.305	-0.689	-0.689 *	-0.578	-0.780	-0.540
	(0.555)	(0.430)	(0.388)	(0.449)	(0.541)	(0.337)
Unemployment benefits (ii)	0.342	0.439	0.542	0.847 **	0.719	0.727 **
	(0.571)	(0.479)	(0.380)	(0.388)	(0.565)	(0.328)
Home size (i)	0.268	0.376 *	-0.071	0.208	0.183	0.192
	(0.225)	(0.207)	(0.192)	(0.228)	(0.252)	(0.163)
Home size (ii)	0.197	0.322	0.086	0.140	0.059	0.220
	(0.269)	(0.207)	(0.171)	(0.192)	(0.271)	(0.143)
Married (i)	-35.504	-21.928	-2.066	-6.983	17.387	-22.931
	(23.669)	(19.285)	(16.372)	(19.568)	(25.276)	(15.421)
Married (ii)	10.370	-0.958	-12.085	-20.238	10.430	-3.058
	(26.864)	(21.047)	(16.674)	(19.218)	(27.180)	(14.911)
Student (i)	48.297	38.346 *	23.560	4.320	-2.628	23.562
	(30.253)	(22.501)	(20.822)	(25.053)	(35.786)	(18.565)
Student (ii)	21.763	16.683	-7.744	-12.135	-3.007	6.474
	(34.996)	(26.086)	(22.569)	(23.948)	(31.045)	(19.320)
Weight (i)	-2.419	-6.812	-7.124	-22.844 **	-25.814 *	-11.814
	(16.352)	(10.952)	(9.852)	(11.132)	(13.401)	(8.255)
Weight (ii)	11.216	22.988 **	21.388 **	26.031 ***	23.605 **	18.239 **
	(14.437)	(10.603)	(9.404)	(9.048)	(11.045)	(7.457)
$Weight^2$ (i)	-0.010	0.004	0.018	0.140 *	0.153 *	0.050
	(0.112)	(0.075)	(0.069)	(0.078)	(0.092)	(0.057)
$Weight^2$ (ii)	-0.029	-0.109	-0.079	-0.113 *	-0.110	-0.063
	(0.100)	(0.073)	(0.065)	(0.062)	(0.072)	(0.052)
Age (i)	34.032	-0.010	-17.079	-19.052	-20.056	-0.984
	(57.090)	(43.502)	(32.998)	(38.446)	(50.820)	(31.604)
Age (ii)	26.236	30.125	77.768 **	58.312	80.949	65.563 *
	(63.144)	(44.987)	(34.394)	(38.459)	(55.363)	(33.800)
Age^2 (i)	-0.679	-0.107	0.193	0.326	0.362	-0.051
	(0.998)	(0.764)	(0.593)	(0.679)	(0.920)	(0.560)
Age^2 (ii)	-0.514	-0.461	-1.216 **	-0.947	-1.311	-1.091 **
	(1.038)	(0.725)	(0.561)	(0.636)	(0.927)	(0.551)
Male child (i)	-18.425	-32.615 **	-25.624 *	-33.839 **	-13.955	-30.625 ***
	(19.717)	(14.770)	(13.300)	(14.540)	(19.145)	(11.199)
Male child (ii)	11.833	-5.623	-9.206	-14.778	-51.292 ***	-5.882
	(19.609)	(15.003)	(12.137)	(14.254)	(18.668)	(11.647)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 11: Estimation results from the AD model using the balanced dataset. CRE added variables.

	Quantile Regressions					
	10%	25%	50%	75%	90%	OLS
Smoked during	-190.485 ***	-112.107 ***	-75.991 ***	-90.337 ***	-2.081	-94.897 ***
	(49.047)	(35.800)	(27.952)	(34.573)	(49.101)	(23.618)
Smoked before	-6.101	-14.429	-30.103	-36.834	-47.618	-21.863
	(35.629)	(27.941)	(24.240)	(26.244)	(37.643)	(19.926)
Drink	11.578	-47.183	-33.420	-45.828	-13.504	-42.091
	(51.397)	(37.665)	(34.045)	(36.649)	(49.262)	(26.156)
Birth control pills	-27.853	-52.478 ***	-28.928 **	-21.657	-19.116	-33.813 ***
	(25.097)	(18.130)	(14.118)	(16.678)	(22.708)	(11.650)
Complications	-122.386 ***	-67.361 ***	-46.073 ***	-29.562 **	-11.029	-65.723 ***
	(24.632)	(16.595)	(12.271)	(14.727)	(21.520)	(10.636)
Doctor visits	5.378	10.838	15.874 **	11.317	-0.099	12.845 *
	(10.409)	(8.081)	(6.916)	(7.094)	(10.166)	(6.765)
Prenatal visits	109.582 ***	90.162 ***	77.477 ***	80.245 ***	73.484 ***	64.011 ***
	(11.583)	(5.906)	(4.583)	(5.080)	(6.969)	(11.670)
Test tube baby	-2.683	64.214	30.674	-57.277	-251.771 ***	-34.417
	(108.153)	(70.387)	(61.737)	(63.434)	(86.622)	(46.240)
Diabetes	181.554 *	217.704 ***	280.896 ***	315.398 ***	373.631 ***	282.529 ***
	(94.004)	(70.214)	(55.456)	(62.333)	(88.341)	(54.962)
Income	-0.126	-0.126	-0.088	-0.127	-0.120	-0.108
	(0.276)	(0.196)	(0.183)	(0.203)	(0.284)	(0.135)
Unemployment benefits	-0.303	-0.185	-0.135	-0.468	0.128	-0.410 *
	(0.510)	(0.358)	(0.311)	(0.350)	(0.484)	(0.239)
Home size	-0.219	0.016	-0.010	0.042	-0.079	-0.072
	(0.268)	(0.191)	(0.180)	(0.225)	(0.294)	(0.143)
Married	-5.950	-1.556	-2.479	-6.801	-19.225	-4.313
	(27.684)	(19.025)	(16.058)	(19.336)	(27.507)	(13.429)
Student	20.743	4.311	1.137	10.260	51.933 *	17.787
	(31.961)	(22.674)	(21.060)	(22.319)	(29.430)	(15.434)
Height	6.662 ***	8.423 ***	10.489 ***	11.419 ***	11.183 ***	10.269 ***
	(1.525)	(1.088)	(0.975)	(1.074)	(1.320)	(0.906)
Weight	5.722	2.412	-8.622	2.040	23.161 **	7.950
	(14.546)	(9.823)	(8.403)	(8.182)	(11.050)	(6.514)
Weight^2	-0.026	0.008	0.063	-0.010	-0.127 *	-0.036
	(0.099)	(0.066)	(0.058)	(0.054)	(0.071)	(0.043)
Age	-40.769	-24.258	-19.409	-29.653 *	-0.204	-14.971
	(30.152)	(20.280)	(16.161)	(17.768)	(26.393)	(14.840)
Age^2	0.867 *	0.516	0.402	0.465	0.065	0.333
	(0.467)	(0.319)	(0.256)	(0.285)	(0.420)	(0.236)
Second child	177.725 ***	162.143 ***	150.112 ***	175.701 ***	156.739 ***	162.162 ***
	(21.562)	(14.632)	(12.540)	(14.267)	(19.457)	(11.515)
Third child	213.334 ***	191.121 ***	190.076 ***	252.706 ***	231.066 ***	209.153 ***
	(33.522)	(24.012)	(22.312)	(24.301)	(31.575)	(20.485)
Fourth child	172.238 ***	182.380 ***	181.915 ***	233.246 ***	202.214 ***	191.631 ***
	(56.800)	(41.156)	(37.683)	(40.316)	(51.240)	(34.217)
Male child	121.921 ***	128.415 ***	137.299 ***	162.926 ***	186.615 ***	145.904 ***
	(16.220)	(12.039)	(10.115)	(12.611)	(16.705)	(8.042)
Education Cat. 1	-17.762	-76.618	-81.724	-39.997	-9.637	-62.830
	(90.017)	(63.440)	(54.685)	(51.733)	(62.760)	(44.738)
Education Cat. 2	40.745	-39.634	-34.390	17.789	63.845	-5.916
	(86.174)	(60.040)	(51.815)	(50.007)	(59.730)	(44.813)
Education Cat. 3	84.970	-8.360	-37.973	-6.862	31.144	-7.380
	(92.000)	(63.567)	(56.402)	(55.421)	(64.993)	(46.993)
Education Cat. 4	57.207	-29.640	-18.659	30.217	85.484	9.992
	(85.169)	(59.404)	(52.030)	(50.103)	(61.687)	(44.595)
Education Cat. 5	67.216	-30.498	-59.437	9.144	77.633	-2.762
	(93.874)	(65.271)	(55.834)	(59.343)	(69.266)	(49.610)
Education Cat. 6	80.965	1.085	-12.183	19.262	66.071	22.282
	(86.395)	(62.012)	(55.002)	(53.307)	(67.144)	(46.642)
Education Cat. 7	56.158	-22.984	-48.886	3.971	40.766	-18.372
	(157.415)	(96.129)	(87.963)	(96.698)	(147.423)	(82.569)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 12: Estimation results from the CREM model using the unbalanced dataset. Main variables.

	Quantile Regressions					
	10%	25%	50%	75%	90%	OLS
Smoked during	1.511 (60.811)	-81.205 * (47.264)	-118.446 *** (36.936)	-118.538 *** (44.050)	-224.515 *** (57.950)	-123.482 *** (30.579)
Smoked before	-2.960 (43.834)	36.036 (34.982)	31.171 (30.794)	63.577 * (33.820)	88.842 * (47.633)	45.941 * (26.418)
Drink	-41.244 (67.546)	12.411 (58.164)	37.354 (45.985)	-0.066 (52.983)	-56.913 (70.383)	18.728 (37.602)
Birth control pills	20.963 (34.027)	42.465 (26.099)	3.423 (22.278)	-5.526 (25.759)	-5.043 (34.068)	13.155 (18.917)
Complications	-135.696 *** (35.896)	-75.278 *** (24.149)	-38.337 * (20.843)	-44.836 * (24.062)	-34.767 (31.528)	-79.684 *** (17.880)
Doctor visits	10.269 (16.818)	1.704 (11.226)	-3.187 (10.423)	-1.286 (12.417)	0.357 (17.171)	5.962 (8.744)
Prenatal visits	8.670 (8.377)	18.729 ** (8.150)	31.123 *** (7.073)	28.191 *** (7.784)	19.459 ** (9.857)	20.307 *** (6.213)
Test tube baby	9.859 (136.485)	-63.569 (97.028)	-5.916 (80.492)	101.877 (97.485)	312.990 ** (129.928)	53.572 (69.107)
Income	0.327 (0.315)	0.248 (0.252)	0.106 (0.225)	0.100 (0.255)	0.018 (0.354)	0.134 (0.175)
Unemployment benefits	0.503 (0.711)	0.260 (0.542)	0.591 (0.498)	0.938 * (0.556)	0.165 (0.743)	0.791 * (0.404)
Home size	0.504 (0.317)	0.262 (0.250)	0.253 (0.220)	0.223 (0.276)	0.272 (0.342)	0.322 * (0.193)
Married	2.626 (32.593)	-11.628 (22.949)	-5.422 (19.948)	13.333 (22.826)	36.962 (31.406)	1.738 (16.984)
Student	42.309 (40.198)	50.671 (31.235)	58.122 ** (27.134)	36.587 (29.995)	-10.418 (40.104)	35.052 (22.354)
Weight	18.993 (14.521)	25.524 ** (10.053)	33.761 *** (9.554)	18.262 ** (8.925)	-0.876 (12.387)	16.575 ** (7.507)
Weight^2	-0.106 (0.098)	-0.158 ** (0.067)	-0.188 *** (0.067)	-0.078 (0.060)	0.035 (0.080)	-0.084 * (0.050)
Age	52.479 (39.830)	33.363 (27.214)	20.231 (21.470)	7.749 (24.502)	5.752 (34.454)	16.427 (20.514)
Age^2	-1.115 * (0.644)	-0.702 (0.438)	-0.460 (0.352)	-0.112 (0.398)	-0.163 (0.550)	-0.382 (0.332)
Male child	-15.270 (25.973)	-30.872 (20.708)	-34.649 * (18.213)	-41.648 ** (19.345)	-46.620 * (24.349)	-30.403 ** (15.446)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 13: Estimation results from the CREM model using the unbalanced dataset. CRE added variables.

	Quantile Regressions					
	10%	25%	50%	75%	90%	OLS
Smoked during	-70.869 ***	-83.486 ***	-99.452 ***	-104.189 ***	-108.669 ***	-94.360 ***
	(26.305)	(21.007)	(19.840)	(20.379)	(24.320)	(19.320)
Smoked before	-55.565 ***	-41.376 ***	-19.328	-10.835	-3.529	-24.233 *
	(18.267)	(15.541)	(14.587)	(15.951)	(18.783)	(14.206)
Drink	-24.374	-22.919	-44.797 *	-51.251 **	-59.697 *	-39.255 *
	(29.747)	(24.710)	(25.253)	(25.508)	(30.727)	(23.552)
Birth control pills	-36.884 **	-31.089 **	-37.452 ***	-30.736 **	-27.382 *	-34.084 ***
	(14.850)	(12.219)	(11.802)	(12.597)	(14.772)	(11.521)
Complications	-96.524 ***	-74.428 ***	-54.155 ***	-31.219 ***	-22.950 *	-64.428 ***
	(16.128)	(12.741)	(11.166)	(11.377)	(13.732)	(11.104)
Doctor visits	14.708 *	10.948	12.629 *	13.281 **	9.717	12.974 *
	(8.338)	(7.224)	(6.623)	(6.666)	(6.830)	(7.323)
Prenatal visits	76.715 ***	70.372 ***	69.121 ***	69.656 ***	68.376 ***	62.657 ***
	(14.315)	(10.600)	(9.265)	(8.646)	(8.266)	(13.404)
Test tube baby	-45.804	-53.412	-18.727	-50.580	-39.651	-35.884
	(56.989)	(47.253)	(39.209)	(40.931)	(48.205)	(40.673)
Income	-0.094	-0.095	-0.114	-0.136	-0.135	-0.109
	(0.147)	(0.124)	(0.119)	(0.127)	(0.152)	(0.116)
Unemployment benefits	-0.311	-0.259	-0.411	-0.231	-0.247	-0.365
	(0.347)	(0.275)	(0.258)	(0.270)	(0.326)	(0.255)
Home size	-0.043	-0.128	-0.098	-0.007	-0.120	-0.086
	(0.145)	(0.120)	(0.115)	(0.122)	(0.138)	(0.112)
Married	-1.042	-1.205	-1.521	-4.776	-4.244	-1.310
	(13.302)	(11.668)	(10.787)	(11.433)	(13.451)	(10.709)
Student	37.874 **	14.307	5.707	3.321	-16.208	5.924
	(17.172)	(14.461)	(13.833)	(14.664)	(17.734)	(13.347)
Weight	6.290	9.372 **	7.402 **	5.478	10.136 **	7.808 **
	(4.425)	(3.942)	(3.468)	(3.594)	(4.093)	(3.539)
Weight^2	-0.036	-0.052 *	-0.033	-0.015	-0.042	-0.036
	(0.031)	(0.027)	(0.024)	(0.025)	(0.028)	(0.024)
Age	-9.353	-20.214	-24.367 **	-28.692 **	-40.657 ***	-23.597 *
	(16.801)	(13.553)	(12.277)	(12.582)	(14.517)	(12.242)
Age^2	0.009	0.177	0.254	0.350 *	0.541 **	0.245
	(0.278)	(0.228)	(0.206)	(0.210)	(0.241)	(0.204)
Second child	153.467 ***	156.642 ***	159.453 ***	148.996 ***	150.982 ***	154.368 ***
	(13.608)	(11.138)	(10.753)	(10.909)	(13.908)	(10.496)
Third child	173.195 ***	184.148 ***	196.344 ***	194.057 ***	201.286 ***	187.332 ***
	(23.305)	(19.337)	(18.227)	(19.007)	(23.195)	(18.328)
Fourth child	183.072 ***	207.923 ***	206.771 ***	203.407 ***	208.761 ***	194.235 ***
	(45.266)	(39.430)	(35.272)	(36.674)	(44.848)	(35.431)
Male child	144.927 ***	140.211 ***	142.491 ***	149.684 ***	143.478 ***	145.187 ***
	(11.433)	(9.381)	(8.734)	(9.158)	(11.089)	(8.548)
Education Cat. 1	-38.800	-25.320	27.994	50.535	104.376 **	21.033
	(57.321)	(33.010)	(27.240)	(35.568)	(50.233)	(27.652)
Education Cat. 2	-49.741	-29.418	33.218	57.239 *	110.142 **	22.576
	(54.799)	(32.347)	(27.109)	(34.744)	(48.344)	(26.771)
Education Cat. 3	-57.021	-29.995	31.191	44.736	105.960 **	18.574
	(58.305)	(35.817)	(29.071)	(37.073)	(51.821)	(28.921)
Education Cat. 4	-48.247	-21.351	33.079	52.249	112.424 **	22.250
	(56.191)	(33.309)	(27.153)	(34.958)	(48.894)	(26.871)
Education Cat. 5	-56.763	-17.026	37.185	60.919	113.416 **	25.409
	(57.618)	(34.794)	(29.122)	(37.192)	(53.932)	(28.824)
Education Cat. 6	-41.767	-26.221	36.781	50.065	104.212 **	21.113
	(57.214)	(33.233)	(27.344)	(35.580)	(50.290)	(27.318)
Education Cat. 7	-109.022	-46.016	29.793	34.105	112.564	7.597
	(90.562)	(61.026)	(48.152)	(65.498)	(99.115)	(48.731)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 14: Estimation results from the 2SFE model using the unbalanced dataset.

	Quantile Regressions				
	10%	25%	50%	75%	90%
Smoked during	-162.837 ***	-156.775 ***	-146.782 ***	-186.265 ***	-183.502 ***
	(34.303)	(25.604)	(23.264)	(24.305)	(32.937)
Smoked before	-24.508	-14.232	-5.846	12.113	29.442
	(26.155)	(20.022)	(17.802)	(19.141)	(27.079)
Drink	-41.675	-33.005	-49.502 *	-52.400 *	-43.625
	(38.308)	(29.008)	(28.939)	(29.333)	(38.988)
Birth control pills	-22.839	-22.815	-24.609 *	-7.255	-26.680
	(19.812)	(13.959)	(13.109)	(13.852)	(19.669)
Complications	-151.500 ***	-93.982 ***	-62.792 ***	-24.545 *	-8.647
	(20.297)	(13.777)	(12.187)	(13.502)	(19.534)
Doctor visits	7.704	8.782	14.639 **	13.681 **	8.451
	(9.085)	(7.003)	(6.103)	(6.370)	(8.415)
Prenatal visits	101.149 ***	88.274 ***	82.826 ***	82.325 ***	82.236 ***
	(12.888)	(7.595)	(5.724)	(5.359)	(6.441)
Test tube baby	43.098	21.891	33.384	5.572	-35.821
	(68.260)	(49.740)	(38.826)	(44.940)	(61.586)
Income	0.010	-0.015	-0.043	-0.070	-0.103
	(0.183)	(0.142)	(0.136)	(0.153)	(0.193)
Unemployment benefits	-0.268	-0.277	-0.142	0.142	-0.357
	(0.411)	(0.303)	(0.259)	(0.305)	(0.390)
Home size	0.094	0.157	0.274 **	0.312 **	0.318 *
	(0.163)	(0.142)	(0.136)	(0.137)	(0.178)
Married	-7.534	2.276	-1.178	-3.597	-13.208
	(17.240)	(13.761)	(13.302)	(14.068)	(18.967)
Student	86.441 ***	63.563 ***	45.732 ***	24.006	33.635
	(22.890)	(16.876)	(17.352)	(17.373)	(25.137)
Weight	31.928 ***	27.255 ***	27.080 ***	23.050 ***	26.062 ***
	(5.952)	(5.399)	(4.906)	(4.881)	(5.529)
Weight^2	-0.170 ***	-0.132 ***	-0.127 ***	-0.094 ***	-0.106 ***
	(0.042)	(0.038)	(0.035)	(0.034)	(0.039)
Age	-4.217	-29.836 *	-30.644 **	-32.730 **	-44.297 **
	(19.625)	(15.671)	(13.528)	(14.332)	(19.300)
Age^2	0.058	0.485 *	0.526 **	0.545 **	0.733 **
	(0.325)	(0.259)	(0.222)	(0.235)	(0.318)
Second child	186.303 ***	169.990 ***	161.051 ***	160.180 ***	168.794 ***
	(18.279)	(13.763)	(12.066)	(14.487)	(19.541)
Third child	167.977 ***	202.422 ***	202.170 ***	220.322 ***	230.686 ***
	(30.166)	(24.527)	(22.574)	(24.571)	(31.334)
Fourth child	191.796 ***	183.375 ***	204.031 ***	228.417 ***	191.609 ***
	(55.278)	(46.668)	(40.472)	(42.210)	(54.352)
Male child	146.002 ***	140.324 ***	143.942 ***	152.610 ***	152.521 ***
	(14.798)	(10.770)	(9.261)	(10.380)	(14.746)
Education Cat. 1	-171.625 *	-69.412	33.134	46.669	-78.829
	(90.106)	(73.051)	(69.667)	(69.450)	(80.900)
Education Cat. 2	-74.140	27.718	128.365 *	148.536 **	41.025
	(87.029)	(69.838)	(66.251)	(67.651)	(80.363)
Education Cat. 3	-31.520	49.709	162.331 **	142.743 *	21.934
	(94.834)	(78.860)	(75.283)	(76.471)	(91.134)
Education Cat. 4	-72.511	47.528	154.641 **	173.598 **	63.779
	(89.993)	(71.321)	(66.672)	(68.443)	(83.084)
Education Cat. 5	-80.900	-1.067	115.933	152.616 **	85.175
	(97.066)	(79.928)	(75.752)	(76.968)	(90.828)
Education Cat. 6	-39.452	71.311	165.983 **	162.931 **	36.387
	(93.207)	(73.391)	(69.194)	(70.773)	(87.325)
Education Cat. 7	-111.444	7.248	183.595 *	172.339	46.874
	(174.631)	(128.867)	(106.956)	(116.731)	(182.347)

Asterisks denote the significance level (double-sided). *: 10%, **: 5%, ***: 1%.
Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 3,000 births and 499 iterations.

Table 15: Estimation results from the post estimated KFE(5) model with $\lambda = 0.8$ and using the unbalanced dataset.

References

Abrevaya J (2001) The Effects of Demographics and Maternal Behavior on the Distribution of Birth Outcomes. Empirical Economics 26(1):247–257

Abrevaya J (2006) Estimating the Effect of Smoking on Birth Outcomes Using a Matched Panel Data Approach. Journal of Applied Econometrics 21(4):489–519

Abrevaya J, Dahl C (2008) The effects of birth inputs on birthweight. Journal of Business and Economic Statistics 26(4):379–397

Almond D, Chay KY, Lee DS (2005) The Costs Of Low Birth Weight. The Quarterly Journal of Economics 120(3):1031–1083

Angrist J, Chernozhukov V, Fernández-Val I (2006) Quantile regression under misspecification, with an application to the us wage structure. Econometrica 74(2):539–563

Arulampalam W, Naylor RA, Smith J (2007) Am I Missing Something? The Effects of Absence From Class on Student Performance. Warwich Economic Research Paper 820, University of Warwick, Department of Economics

Bache SH (2010) Minimax regression quantiles. CREATES Research Paper 2010-54, CREATES, Aarhus University

Bernstein IM, Mongeon JA, Badger GJ, Solomon L, Heil SH, Higgins ST (1978) Maternal Smoking and Its Association With Birth Weight. Obstetrics & Gynecology 106(5):986–991

Bernstein IM, Horbar JD, Badger GJ, Ohlsson A, Golan A (2000) Morbidity and Mortality Among Very-Low-Birth-Weight Neonates with Intrauterine Growth Restriction. American Journal of Obstetrics and Gynecology 182(1):196–206

Black SE, Devereux PJ, Salvanes KG (2007) From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. Quarterly Journal of Economics 122(1):409–439

Buckles K, Hungerman D (2008) Season of birth and later outcomes: Old questions, new answers. NBER Working Paper 14573, National Bureau of Economic Research

Canay I (2010) A simple approach to quantile regression for panel data. Working paper, Northwestern University

Chamberlain G (1984) Panel Data. In: Griliches Z, Intriligator MD (eds) Handbook of Econometrics, vol 2, Elsevier Science B. V., pp 1247–1318

Chernozhukov V (2010) Inference for extremal conditional quantile models, with an application to birthweights. Review of Economic Studies (forthcoming)

Corman H, Chaikind S (1998) The Effect of Low Birthweight on the School Performance and Behavior of School-Aged Children. Economics of Educations Review 17(3):307–316

Dehejia R, Lleras-Muney A (2004) Booms, busts, and babies' health. Quarterly Journal of Economics 119(3):1091–1130

DiFranza JR, Aligne CA, Weitzman M (2004) Prenatal and Postnatal Environmental Tobacco Smoke Exposure and Children's Health. Pediatrics 113(4):1007–1015

Fitzenberger B, Kohn K, Wang Q (2010) The erosion of union membership in germany: determinants, densities, decompositions. Journal of Population Economics 24(1):141–165

Hack M, Klein NK, Taylor HG (1995) Long-term developmental outcomes of low birth weight infants. The Future of Children 5(1):176–196

Hofhuisi W, de Jongste JC, Merkus PJFM (2003) Adverse Health Effects of Prenatal and Postnatal Tobacco Smoke Exposure on Children. Archives of Disease in Childhood 88(12):1086–1090

Kirkegaard I, Obel C, Hedegaard M, Henriksen TB (2006) Gestational Age and Birth Weight in Relation to School Performance of 10-Year-Old Children: A Follow-up Study of Children Born After 32 Completed Weeks. Pediatrics 118(4):1600–1606

Koenker R (2004) Quantile regression for longitudinal data. Journal of Multivariate Analysis 91(1):74–89

Koenker R, Hallock KF (2001) Quantile Regression. The Journal of Economic Perspectives 15(4):143–156

Linnet KM, Obel C, Bonde E, Hove P, Thomsen, Secher NJ, Wisborg K, Henriksen TB (2006) Cigarette Smoking During Pregnancy and Hyperactive-Distractible Preschooler's: A follow-up Study. Acta Paediatrica 95(6):694–700

Mundlak Y (1978) On the Pooling of Time Series and Cross Section Data. Econometrica 46(1):69–85

Permutt T, Hebel JR (1989) Simultaneous-Equation Estimation in a Clinical Trial of the Effect of Smoking on Birth Weight. Biometrics 45(2):619–622

Royer H (2004) What All Women (and Some Men) Want to Know: Does Maternal Age Affect Infant Health? Working Paper 68, Center of Labor Economics, University of California, Berkely

Wang X, Zuckerman B, Pearson C, Kaufman G, Chen C, Wang G, Niu T, Wise PH, Bauchner H, Xu X (2002) Maternal Cigarette Smoking, Metabolic Gene Polymorphism, and Infant Birth Weight. The Journal of the American Medical Association 287(2):195–202

Wisborg K, Kesmodel U, Henriksen TB, Olsen SF, Secher NJ (2000) A Prospective Study of Smoking During Pregnancy and SIDS. Archives of Disease in Childhood 83(3):203–206

Wisborg K, Kesmodel U, Henriksen TB, Olsen SF, Secher NJ (2001) Exposure to Tobacco Smoke in Utero and the Risk of Stillbirth and Death in the First Year of Life. American Journal of Epidemiology 154(4):322–327

Research Papers 2008

2008-07	Changli He, Annastiina Silvennoinen and Timo Teräsvirta: Parameterizing unconditional skewness in models for financial time series
2008-08	Cristina Amado and Timo Teräsvirta: Modelling Conditional and Unconditional Heteroskedasticity with Smoothly Time-Varying Structure
2008-09	Søren Johansen and Bent Nielsen: An analysis of the indicator saturation estimator as a robust regression estimator
2008-10	Peter Christoffersen, Kris Jacobs, Christian Dorion and Yintian Wang: Volatility Components, Affine Restrictions and Non-Normal Innovations
2008-11	Peter Christoffersen, Kris Jacobs, Chayawat Ornthanalai and Yintian Wang: Option Valuation with Long-run and Short-run Volatility Components
2008-12	Tom Engsted and Stig V. Møller: An iterated GMM procedure for estimating the Campbell-Cochrane habit formation model, with an application to Danish stock and bond returns
2008-13	Lars Stentoft: Option Pricing using Realized Volatility
2008-14	Jie Zhu: Pricing Volatility of Stock Returns with Volatile and Persistent Components
2008-15	Jie Zhu: Testing for Expected Return and Market Price of Risk in Chinese A-B Share Market: A Geometric Brownian Motion and Multivariate GARCH Model Approach
2008-16:	Jie Zhu: FIEGARCH-M and and International Crises: A Cross-Country Analysis
2008-17:	Almut E. D. Veraart: Inference for the jump part of quadratic variation of Itô semimartingales
2008-18:	Michael Sørensen: Parametric inference for discretely sampled stochastic differential equations
2008-19:	Anne Péguin-Feissolle, Birgit Strikholm and Timo Teräsvirta: Testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form
2008-20:	Stefan Holst Bache, Christian M. Dahl and Johannes Tang Kristensen: Determinants of Birthweight Outcomes: Headlights on tobacco road to low birthweight outcomes - Evidence from a battery of quantile regression estimators and a heterogeneous panel