# The effects of smoking and prenatal care on birth outcomes: evidence from quantile estimation on panel data

by

Jason Abrevaya and Christian M. Dahl

Department of Economics
Purdue University

# Motivation

- Adverse birth outcomes have been found to result in large economic costs, in the form of both direct medical costs and long-term developmental consequences.
  - Almond et. al. (2004) report that the hospital costs for newborns peaks at around $150,000 (in 2000 dollars) for infants that weigh 800 grams; the costs remain quite high for all "low birthweight" outcomes, with an average cost of around $15,000 for infants that weigh 2000 grams.
  - The infant-mortality rate increases at lower birthweights.

- LBW babies ($< 2500$ gram) have developmental problems in cognition, attention, and neuromotor functioning that persist until adolescence (Hack et. al. (1995)).

- LBW babies are more likely to delay entry into kindergarten, repeat a grade in school, and attend special-education classes (Corman (1995); Corman and Chaikind (1998)).

- LBW babies are more likely to have inferior labor-market outcomes, being more likely to be unemployed and earn lower wages (Behrman and Rosenzweig (2004); Case et. al. (2005); Currie and Hyson (1999)).

- An enormous difficulty in evaluating initiatives aimed at improving birth outcomes is to accurately estimate the causal effects of prenatal activities on these birth outcomes.

- Unobserved heterogeneity among childbearing women makes it difficult to isolate the causal effects of smoking and prenatal care on birth outcomes (such as birthweight). Whether or not a mother smokes, for instance, is likely to be correlated with unobserved characteristics of the mother.

- To deal with this difficulty, various studies have used an instrumental-variable methodology in order to estimate the effects of smoking (Evans and Ringel (1999); Permutt and Hebel (1989)) and prenatal care (Currie and Gruber (1996); Evans and Lien (2005); Joyce (1999)) on birth outcomes.

- Another approach has been to utilize panel data (i.e., several births for each mother) in order to identify these effects from changes in prenatal behavior or maternal characteristics between pregnancies (Abrevaya (2005); Currie and Moretti (2002); Rosenzweig and Wolpin (1991); Royer (2004)).

- A potential limitation of these studies is that they have focused upon estimating how prenatal behaviors and maternal characteristics affect *average* birthweight.

- In contrast, the costs associated with birthweight have been found to exist primarily at the low end of the birthweight distribution, with the costs increasing significantly at the very low end.

- Abrevaya (2001) uses cross-sectional federal natality data and finds that various observables (such as smoking) have significantly larger effects at lower quantiles of the birthweight distribution (but does not account for unobserved heterogeneity).

- This paper focuses upon estimation of the prenatal effects on the entire birthweight distribution. Using state-level (maternally linked) panel data on births to control for unobserved heterogeneity, we address a major shortcoming of previous work on the association between prenatal care and the birthweight distribution.

- This paper combines the panel-data (correlated random effects model) methodology of previous studies with a focus upon estimation of effects on the quantiles of birthweight. In particular, we focus upon a notion of marginal effects upon conditional quantiles that is analogous to the standard notion of marginal effects upon the conditional expectation. These effects explicitly control for unobserved heterogeneity by allowing the "mother random effect" to be related to observables.

- This approach is an important methodological contribution to the literature, as it provides a general framework with which empirical researchers can apply quantile regression to panel data.

- We find some interesting differences between the panel-data and cross-sectional results. For example, the results from panel-data estimation, which controls for unobserved heterogeneity, indicate that the negative effects of smoking on birthweight are significantly lower (in magnitude) across all quantiles than indicated by the cross-sectional estimates.

# The idea behind quantile regression

- In the traditional least-squares regression the conditional mean function that describes how the mean of $y$ changes with the vector of covariates $\boldsymbol{x}$, is (almost) all we need to know about the relationship between $y$ and $\boldsymbol{x}$.

- It is assumed that $\boldsymbol{x}$ affects only the location of the conditional distribution of $y$, not its scale, or any other aspects of its distributional shape.

- Is there more to econometric life than is dreamt of in this philosofy of the location shift model?

- Yes? Covariates may influence the conditional distribution of the response in myriad other ways
  - expanding its dispersion (as in traditional models of heterosked.),
  - stretching one tail of the distribution, compressing the other tail,
  - even inducing multimodality.

- Explicit investigation of these effects by looking at the quantiles of the distribution of $y|x$ (and not only the conditional mean) may give a more informative empirical analysis.

# Mathematical definition of quantile regression

- Recall that we can obtain an estimate of $\mathsf{E}(y_i|\boldsymbol{x}_{i\cdot}) = \boldsymbol{x}_{i\cdot}\boldsymbol{\beta}$ by solving

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^k}{\arg\min} \sum_{i=1}^{n} (y_i - \boldsymbol{x}_{i\cdot}\boldsymbol{\beta})^2$$

- Let $\tau \in [0,1]$ denote quantiles and assume that the conditional quantile function is linear in $\boldsymbol{x}$, i.e.,

$$Q_\tau\,(y_i|\boldsymbol{x}_{i\cdot}) = \boldsymbol{x}_{i\cdot}\boldsymbol{\beta}_\tau$$

- The estimate of the conditional quantile function $Q_\tau(y_i|\boldsymbol{x}_{i\cdot})$ is found by solving

$$\widehat{\boldsymbol{\beta}}_\tau = \underset{\boldsymbol{\beta}\in\mathbb{R}^k}{\arg\min}\left[\sum_{i\in\{i:y_i\geq\boldsymbol{x}_{i\cdot}\boldsymbol{\beta}\}}\tau\,|y_i-\boldsymbol{x}_{i\cdot}\boldsymbol{\beta}|+\sum_{i\in\{i:y_i<\boldsymbol{x}_{i\cdot}\boldsymbol{\beta}\}}(1-\tau)\,|y_i-\boldsymbol{x}_{i\cdot}\boldsymbol{\beta}|\right]$$

- Consider the median $\tau = 0.5$. The estimator of $Q_{\tau=0.5}(y_i|\boldsymbol{x}_{i\cdot})$ can then be found by finding the minimizer of the sum of absolute residuals (the LAD estimator)

$$\widehat{\boldsymbol{\beta}}_{\tau=0.5} = \underset{\boldsymbol{\beta}\in\mathbb{R}^k}{\arg\min}\sum_{i=1}^{n}|y_i-\boldsymbol{x}_{i\cdot}\boldsymbol{\beta}|$$

# The correlated random-effects panel data model

- The standard linear unobserved effects panel data model

$$y_{mb} = \boldsymbol{x}'_{mb}\boldsymbol{\beta} + c_m + u_{mb}, \;\; m = 1, ..., n; b = 1, 2$$

  where

  - $m$ indexes individuals (mothers) and $b$ indexes time periods (births),
  - $y$ denotes a response and $\boldsymbol{x}$ denotes a vector of observables,
  - $c$ denotes the unobservable individual effects,
  - $u$ denotes an idiosyncratic disturbance.

- What if $c_m$ is correlated with $\boldsymbol{x}'_{mb}$?
  - If ignored $\widehat{\boldsymbol{\beta}}$ will be biased/inconsistent

- Fixed-effects model:
  - Exploit that an expectation is a linear operator, such that

$$\begin{aligned} \mathsf{E}(y_{m2} - y_{m1} | \boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1}) &= \mathsf{E}(y_{m2} | \boldsymbol{x}'_{m2}) - \mathsf{E}(y_{m1} | \boldsymbol{x}'_{m1}) \\ &= (\boldsymbol{x}'_{m2} - \boldsymbol{x}'_{m1})' \boldsymbol{\beta}. \end{aligned}$$

  - Consequently (under regularity conditions) a consistent estimator of $\boldsymbol{\beta}$ can be obtained by regressing $y_{m2} - y_{m1}$ on $\boldsymbol{x}'_{m2} - \boldsymbol{x}'_{m1}$

- Correlated random-effects (Chamberlain, 1982, 1984)
  - Views the unobservable $c_m$ as a linear projection onto the observables plus a disturbance

    $$c_m = \psi + \boldsymbol{x}'_{m1}\boldsymbol{\lambda}_1 + \boldsymbol{x}'_{m2}\boldsymbol{\lambda}_2 + v_m,$$

    where $\psi$ is a scalar and $v$ is a disturbance that (by definition of linear projections) is uncorrelated with $\boldsymbol{x}_{m1}$ and $\boldsymbol{x}_{m2}$
  - Inserting this expression for $c_m$ in the panel data model yields

    $$\begin{aligned} y_{m1} &= \psi + \boldsymbol{x}'_{m1}(\boldsymbol{\beta} + \boldsymbol{\lambda}_1) + \boldsymbol{x}'_{m2}\boldsymbol{\lambda}_2 + v_i + u_{i1} \\ y_{m2} &= \psi + \boldsymbol{x}'_{m1}\boldsymbol{\lambda}_1 + \boldsymbol{x}'_{m2}(\boldsymbol{\beta} + \boldsymbol{\lambda}_2) + v_i + u_{i2}. \end{aligned}$$

  - Consequently (under regularity conditions) a consistent estimator of $\boldsymbol{\beta}$ is typically obtained by the method of Minimum Distance.

# Estimated effects: The meaning of $\boldsymbol{\beta}$

- In the fixed-effects model

$$\boldsymbol{\beta} = \frac{\partial \, \mathsf{E}(y_{m2} - y_{m1} | \boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})}{\partial (\boldsymbol{x}'_{m2} - \boldsymbol{x}'_{m1})'}.$$

- In the correlated random-effects model

$$\boldsymbol{\beta} = \frac{\partial \, \mathsf{E}(y_{m1} | \boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})}{\partial \boldsymbol{x}'_{m1}} - \frac{\partial \, \mathsf{E}(y_{m2} | \boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})}{\partial \boldsymbol{x}'_{m1}} \tag{1}$$

- $\boldsymbol{\beta}$ tells us how much $\boldsymbol{x}_{m1}$ affects $\mathsf{E}(y_{m1} | \boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})$ above and beyond the effect that works through the unobservable $c_m$

# Conditional quantiles with panel data

- Problem in fixed-effect model approach since conditional quantiles in general are not linear operators, i.e,

  $$Q_\tau(y_{m2}-y_{m1}|\boldsymbol{x}'_{m2},\boldsymbol{x}'_{m1}) \neq Q_\tau(y_{m2}|\boldsymbol{x}'_{m2},\boldsymbol{x}'_{m1})-Q_\tau(y_{m1}|\boldsymbol{x}'_{m2},\boldsymbol{x}'_{m1})$$

  so the differencing approach does not seem to be working well for quantile regression panel data models:

  – Consistency and interpretation of $\boldsymbol{\beta}$ is questionable?

- Koenker and Hallock (2000) write: *Quantiles of convolutions of random variables are rather intractable objects, and preliminary differencing strategies familiar from Gaussian models have sometimes unanticipated effects.*

- In quantile regression correlated random effect model the effects of the observables on the response variable for a given quantile can be given analogous to the right hand side of (1) as

$$\frac{\partial Q_\tau(y_{m1}|\boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})}{\partial \boldsymbol{x}'_{m1}} - \frac{\partial Q_\tau(y_{m2}|\boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})}{\partial \boldsymbol{x}'_{m1}} \tag{2}$$

- For example the difference in (2) is the effect of $\boldsymbol{x}'_{m1}$ (first period observables) on $Q_\tau(y_{m1}|\boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})$ above and beyond the effect on the $\tau$-th conditional quantile that works through the unobservable $c_m$.

- To compute the estimated effects we then need a model for both $Q_\tau(y_{m1}|\boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})$ and $Q_\tau(y_{m2}|\boldsymbol{x}'_{m2}, \boldsymbol{x}'_{m1})$

- Crucial modelling assumption

$$
\begin{aligned}
Q_\tau(y_{m1}|\boldsymbol{x}'_{m2},\boldsymbol{x}'_{m1}) &= \phi^1_\tau + \boldsymbol{x}'_{m1}(\boldsymbol{\beta}_\tau + \boldsymbol{\lambda}^1_\tau) + \boldsymbol{x}'_{m2}\boldsymbol{\lambda}^2_\tau && (3)\\
Q_\tau(y_{m2}|\boldsymbol{x}'_{m2},\boldsymbol{x}'_{m1}) &= \phi^2_\tau + \boldsymbol{x}'_{m1}\boldsymbol{\lambda}^1_\tau + \boldsymbol{x}'_{m2}(\boldsymbol{\beta}_\tau + \boldsymbol{\lambda}^2_\tau). && (4)
\end{aligned}
$$

- Consequently we can interpret $\boldsymbol{\beta}$ as the "effects of the observables" since from (3) and (4) we can compute vbeta as

$$
\boldsymbol{\beta} = \frac{\partial Q_\tau(y_{m1}|\boldsymbol{x}'_{m2},\boldsymbol{x}'_{m1})}{\partial \boldsymbol{x}'_{m1}} - \frac{\partial Q_\tau(y_{m2}|\boldsymbol{x}'_{m2},\boldsymbol{x}'_{m1})}{\partial \boldsymbol{x}'_{m1}} \qquad (5)
$$

which obviously equals (2).

# Empirical illustration:

The data: In the United States, detailed “natality data” is recorded for nearly every live birth that occurs. Detailed information on maternal characteristics (age, education, race, etc.), birth outcomes (birthweight, gestation, etc.), and prenatal care (number of prenatal visits, smoking status, etc.) is collected by each state (with federal guidelines on specific data-item requirements).

- *Washington data:* The Washington State Longitudinal Birth Database (WSLBD) was provided by Washington's Center for Health Statistics. The WSLBD is a panel dataset consisting of all births between 1992 and 2002 that could be accurately linked together as belonging to the same mother.

- *Arizona data:* The Arizona Department of Health Services provided the authors with data on all births occurring in the state of Arizona between 1993 and 2002. Although names were not provided, the exact dates of birth for both mother and father were provided in the data.

Table 1: Descriptive Statistics, Washington and Arizona Birth Panels

| Variable | Washington | | Arizona | |
|---|---|---|---|---|
| | 1st Child | 2nd Child | 1st Child | 2nd Child |
| Birthweight (in grams) | 3442 (523) | 3530 (536) | 3339 (517) | 3427 (505) |
| Male child | 0.515 | 0.511 | 0.520 | 0.516 |
| Mother's age | 25.27 (5.25) | 27.89 (5.35) | 25.23 (5.26) | 27.85 (5.36) |
| Mother's education | 13.52 (2.32) | 13.72 (2.21) | 13.21 (2.68) | 13.39 (2.61) |
| Married | 0.751 | 0.853 | 0.780 | 0.886 |
| No prenatal care | 0.004 | 0.003 | 0.005 | 0.006 |
| 1st-trimester care | 0.879 | 0.895 | — | — |
| 2nd-trimester care | 0.107 | 0.093 | — | — |
| 3rd-trimester care | 0.014 | 0.012 | — | — |
| Smoke | 0.143 | 0.132 | 0.049 | 0.044 |
| Drink | 0.017 | 0.014 | 0.009 | 0.007 |
| # prenatal visits | 12.06 (3.53) | 11.63 (3.25) | 11.83 (3.59) | 11.73 (3.55) |
| Year of birth | 1995.0 (2.2) | 1997.8 (2.3) | 1996.3 (2.3) | 1998.9 (2.2) |
| # of Observations | 45,067 | 45,067 | 56,201 | 56,201 |

- Summary
  - The average birthweight increases by 88 grams at the second birth for both Washington mothers and Arizona mothers.
  - For their second birth women are less likely to smoke and drink and more likely to be married, have a male child, and have their first prenatal-care visit during the first trimester.
  - On average, Arizona mothers are slightly less educated, have babies with higher birthweight and have a lower reported rate of drinking during pregnancy.

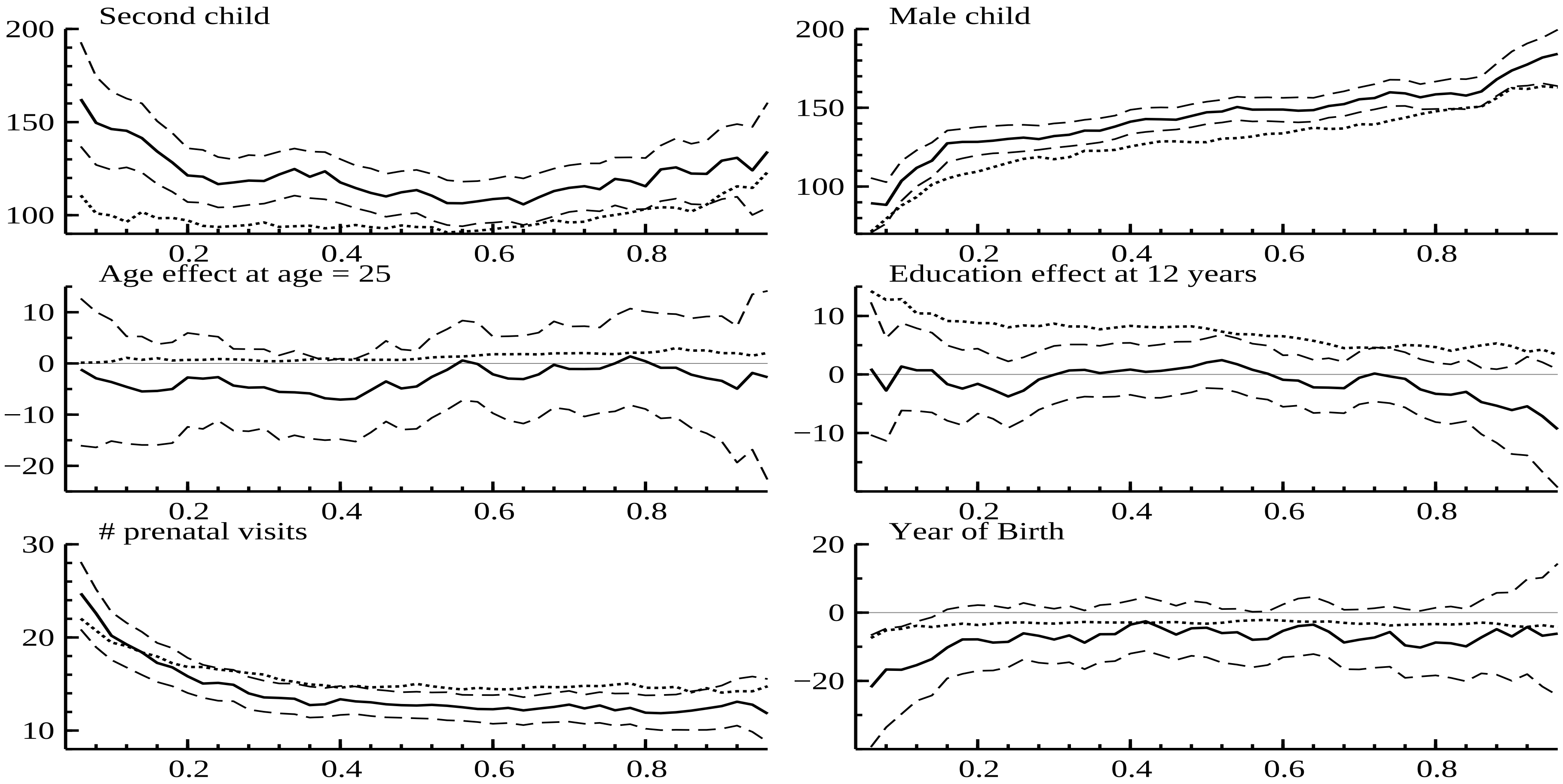
– The largest difference between the two samples appears to be the level of smoking: Washington mothers report smoking in 13.7% of pregnancies (which is right around the national average during this time period), whereas Arizona mothers report smoking in only 4.7% of pregnancies.

Estimation results (Washington):

- *Second child:* Birthweights are uniformly larger for second children at all quantiles, for both the cross-sectional and panel estimates. The panel estimates of the second-child effect are somewhat larger than the cross-sectional estimates, with the largest effects at the lowest quantiles (e.g., 146 grams at the 10% quantile).

- *Male child:* It is well-known that, on average, male babies weigh more at birth than female babies. The quantile estimates indicate that the positive male-child effect on birthweight is present at all quantiles of the conditional birthweight distribution. The magnitude of the effect increases when one moves from lower quantiles to higher quantiles, with the panel estimates indicating a slightly higher effect (10–20 grams) than the cross-sectional estimates.

# Washington

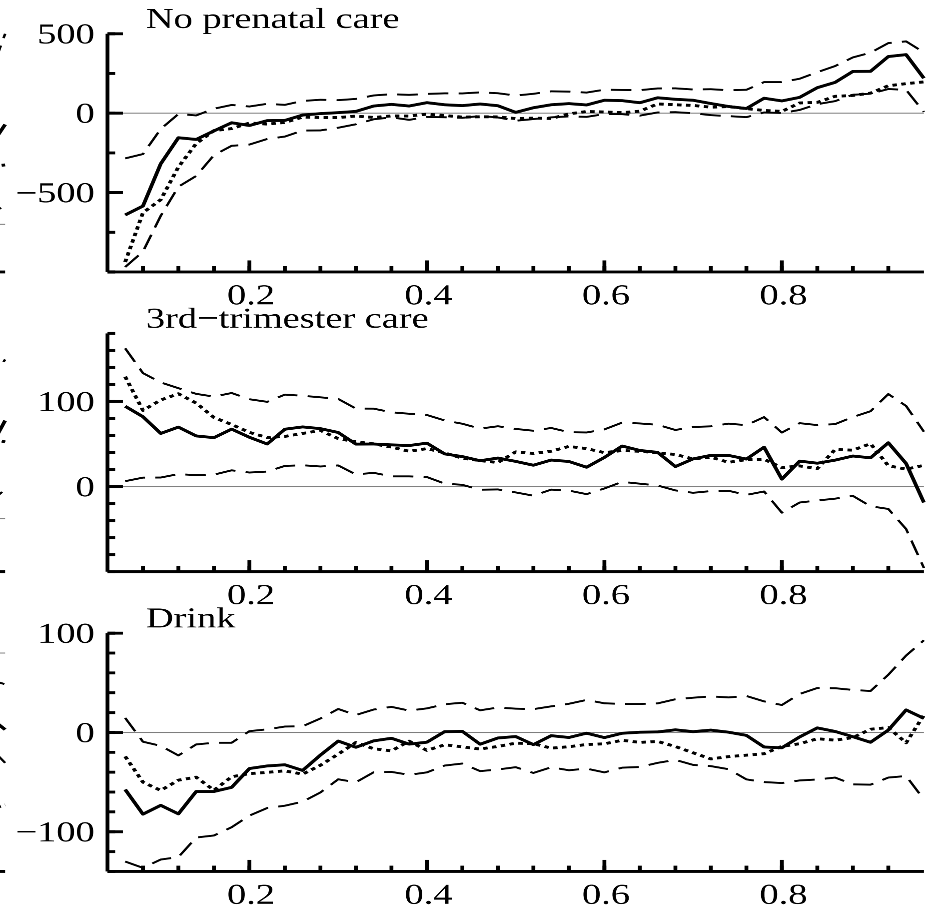

Second child
200
150
100
0.2
0.4
0.6
0.8
Male child
200
150
100
0.2
0.4
0.6
0.8
Age effect at age = 25
10
0
−10
−20
0.2
0.4
0.6
0.8
Education effect at 12 years
10
0
−10
0.2
0.4
0.6
0.8
# prenatal visits
30
20
10
0.2
0.4
0.6
0.8
Year of Birth
20
0
−20
0.2
0.4
0.6
0.8

- *Age and education:* For age, both the cross-sectional and panel estimates are statistically insignificant at a 5% level for all quantiles. For education, the cross-sectional estimates are positive across the quantiles and statistically significant except at quantiles above 80%. In contrast, the panel estimates are statistically insignificant across all quantiles. This difference could be due to two factors:

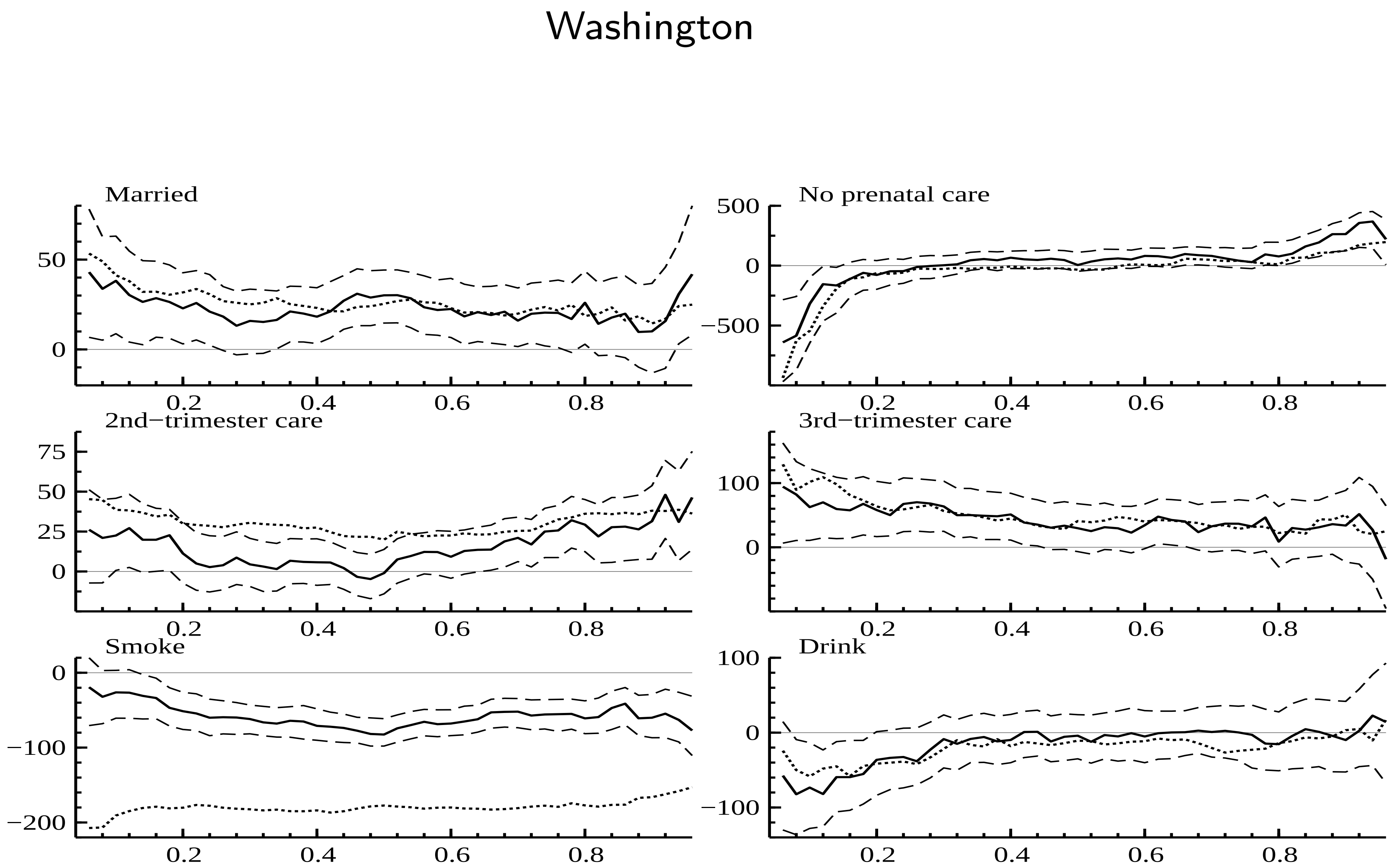
  – the amount of within-mother variation in education is quite small, with the average change in education for the sample being about 0.2 years; and,

  – the level of education may be correlated with the mother-specific unobservable. Years of schooling is likely positively correlated with $c_m$, which would imply that the cross-sectional estimates are biased upwards.

# Washington


Married
50
0
0.2
0.4
0.6
0.8
No prenatal care
500
0
−500
2nd−trimester care
75
50
25
0
3rd−trimester care
100
0
Smoke
0
−100
−200
Drink
100
0
−100

- *Marital status:* The estimated positive effects of marriage on birthweight are quite similar for the cross-sectional and panel specifications, in the 20–50 gram range over the quantiles considered.
  - One should be cautious about interpreting the cross-sectional marriage estimates as causal since marital status is an explanatory variable that *a priore* would appear to serve as a proxy for mother-specific unobservables. The panel estimates are lower than the CS estimates in the lower quantiles, suggesting that this might be a factor in the lower quantiles.
  - The estimates are consistent with a situation in which marriage provides the birth mother with support (support at home, financial support, emotional support, etc.) that would lead to a more favorable birth outcome.


Washington
Married
50
0
0.2
0.4
0.6
0.8
No prenatal care
500
0
−500
2nd−trimester care
75
50
25
3rd−trimester care
100
Smoke
−100
−200
Drink

- *Prenatal-care visits:* It should be pointed out that interpreting the effect of any prenatal-care variable is a bit difficult since the *observed* prenatal care proxies for both *intended* prenatal care and pregnancy problems.

  - For instance, if two mothers have identical intentions (at the beginning of pregnancy) with respect to prenatal-care visits, the mother that experiences problems early in her pregnancy would be more likely to have an earlier first prenatal-care visit and to have more prenatal-care visits overall.

  - The estimated effects of the prenatal-care variables, therefore, may reflect the combined effects of intended care and pregnancy complications.

- The estimates for the no-prenatal-care indicator variable, which are significantly negative at the 10% quantile and significantly *positive* at the 90% quantile, illustrate this point.

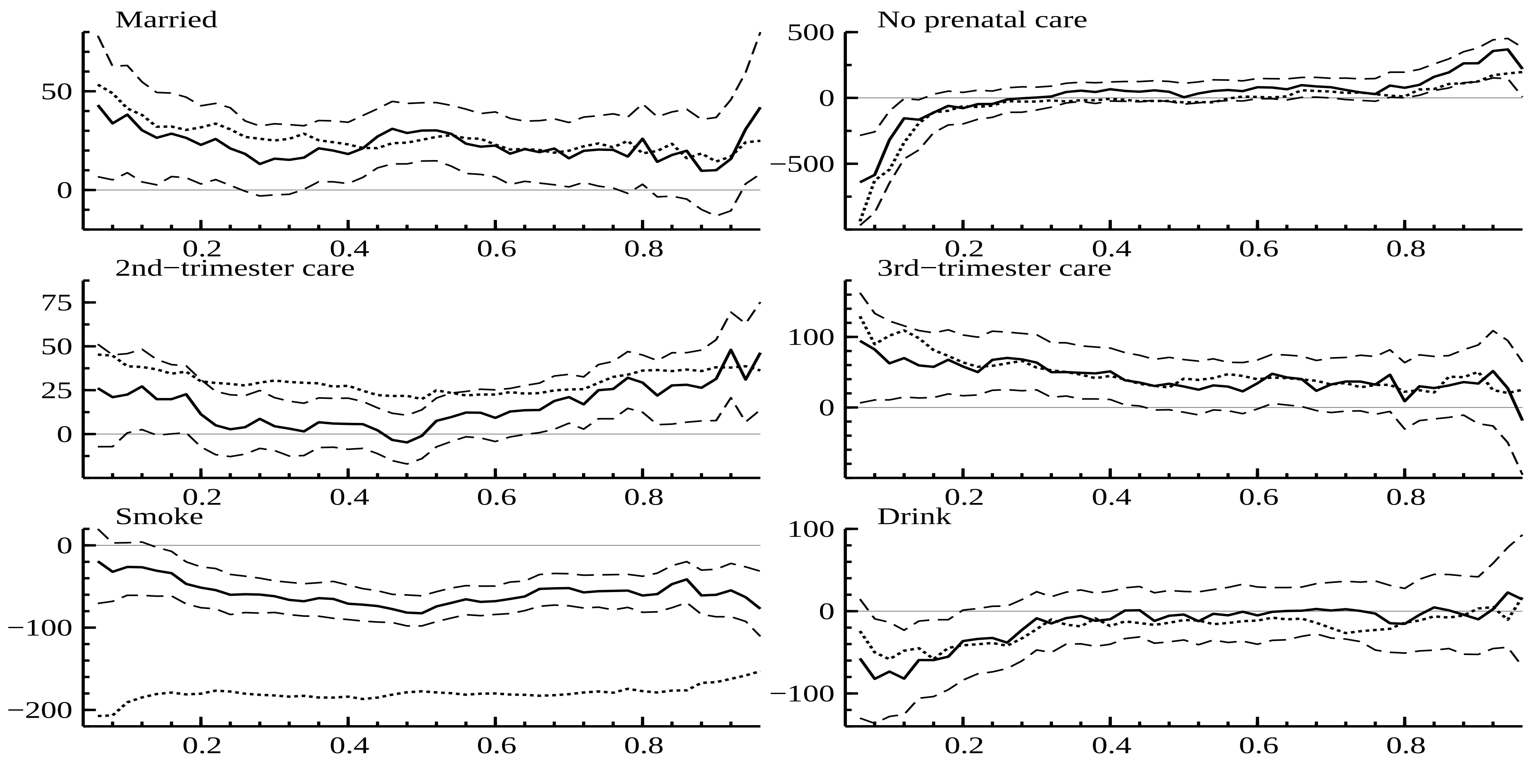
- A possible explanation for the dramatic difference at the two ends of the distribution is that lack of prenatal care is more likely to proxy for lack of intended care at the lowest quantiles and more likely to proxy for a problem-free pregnancy at the highest quantiles. At the intermediate quantiles, the effect of the no-prenatal-care indicator is found to be statistically insignificant in both the cross-sectional and panel results.

- For the third-trimester-care indicator variable, the cross-sectional and panel estimates are similar, indicating positive effects (as compared to first-trimester care) which become less statistically significant at higher quantiles.

- For the indicator variables, the largest difference between the cross-sectional and panel results shows up in the second-trimester-care variable; the cross-sectional estimates are statistically significant at all quantiles and range from 25 to 50 grams, whereas the panel estimates are somewhat lower (close to zero in intermediate quantiles) and only significantly positive at the highest quantiles.

- The effect of the number of prenatal visits is estimated to be significantly positive across all quantiles, with larger effects found at lower quantiles and the effects essentially "flattening out" (at around 14–15 grams per visit for the cross-sectional results and 12–13 grams per visit for the panel results).

# Washington

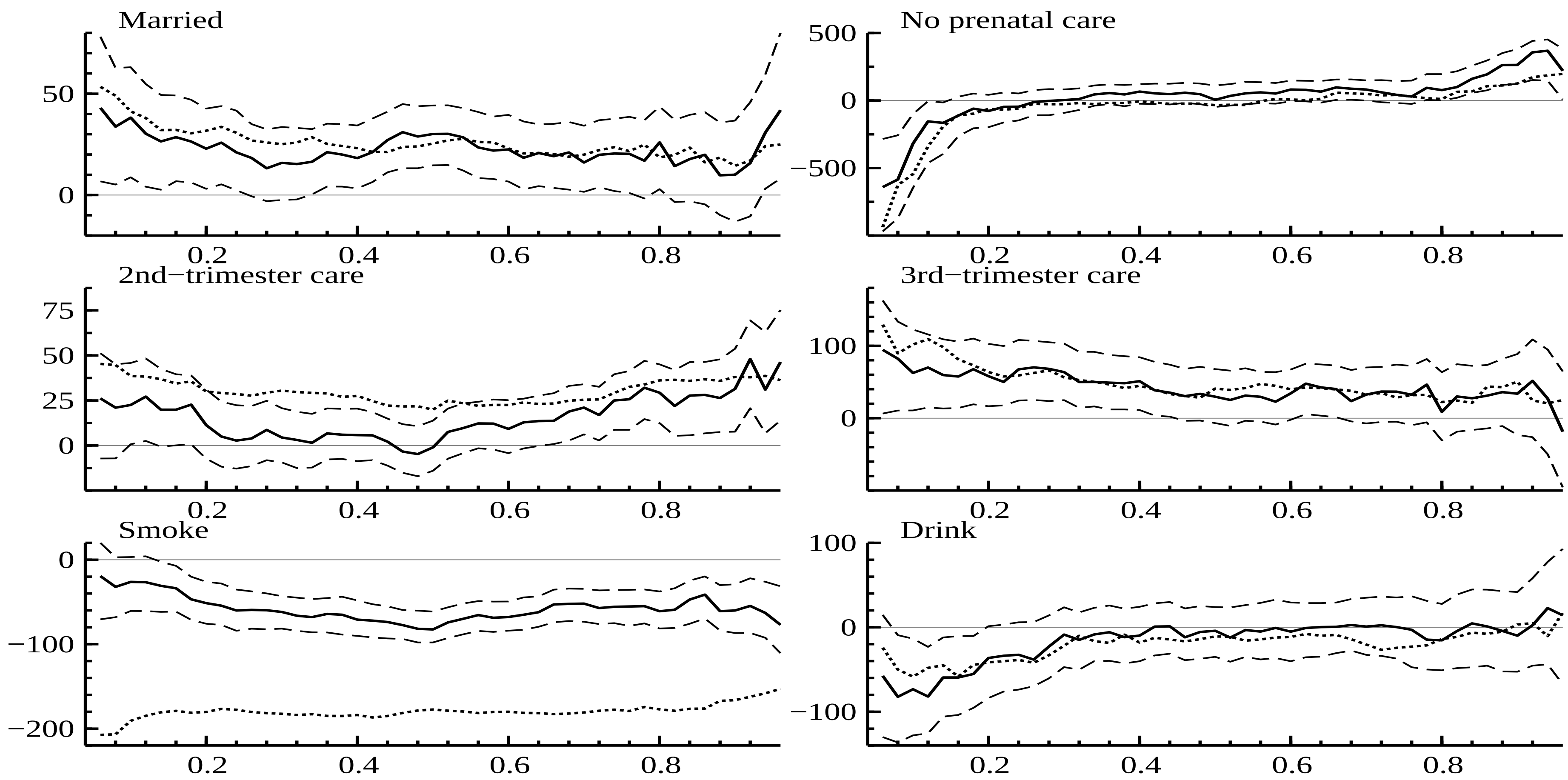


Married
No prenatal care
2nd−trimester care
3rd−trimester care
Smoke
Drink
0.2
0.4
0.6
0.8

- *Smoking:* The most dramatic difference between the cross-sectional and panel results involves the estimated effects of smoking.
  - The cross-sectional results indicate that the negative effects of smoking are in the range of 150–200 grams, with larger effects at lower quantiles. The panel estimates are still significantly negative at all but the lowest quantiles, but the estimated effects are much lower in magnitude (mostly in the 50–80 gram range between the 20% and 80% quantiles).
  - The omitted-variables explanation of this large difference would be that the smoking indicator in the cross-sectional specification is negatively correlated with the error disturbance in the birthweight regression equation.

- Consistent with this explanation, the smoking coefficients in both $\lambda_\tau^1$ and $\lambda_\tau^2$ are found to be significantly negative across the quantiles
- The magnitudes of the cross-sectional and panel estimates are roughly in agreement with those found by Abrevaya (2005) for the (conditional expectation) effects in federal natality data.

# Washington


Married
No prenatal care
2nd−trimester care
3rd−trimester care
Smoke
Drink
0.2
0.4
0.6
0.8
50
0
500
−500
75
25
100
−100
−200

- *Alcohol consumption:* In contrast to the smoking results, the estimated effects of alcohol consumption (as measured by the alcohol-consumption indicator variable) are quite similar for the cross-sectional and panel specifications.
  - Drinking is estimated to have significant negative effects at lower quantiles (below about the 20% quantile), with the magnitudes of the effects ranging between about 40 and 80 grams.
  - Very few mothers actually report that they consumed alcohol during pregnancy (only about 1.5% in our sample).
  - The lack of strong statistical evidence regarding the effects of drinking could stem from the low variation in the indicator variable and the probable large rates of misclassification.

Estimation results (Arizona):

- Overall, there is a remarkable similarity between the results for the two samples.
  - There is a significant positive effect of the second child across all quantiles (50–110 grams from the Arizona panel estimates).
  - The positive birthweight effect of a male child increases from lower to higher quantiles.
  - Despite a positive estimated cross-sectional effect of education at lower quantiles, the panel estimates indicate no significant education effect.

  - The effect of the number of prenatal visits is highest at lower quantiles, with the effect flattening out at higher quantiles. For both Washington and Arizona, the cross-sectional estimate of the effect is lower at lower quantiles and higher at higher quantiles.

  - The magnitude of the negative smoking effect is significantly lower for the panel estimates (ranging between 40 and 80 grams for Arizona) than for the cross-sectional estimates.

- Some differences:

  - Although the cross-sectional estimates of the marriage effect are still significantly positive, the panel-data estimates indicate no statistically significant effect of marriage for Arizona mothers (data matching issues?)

– Drinking is not found to have a statistically significant effect at any of the quantiles (either in the cross section or the panel).

– Due to the lack of indicator variables for second-trimester and third-trimester care, the estimated effects of the no-care indicator variable and the number of prenatal visits are slightly different. The magnitude of the quantile effects for number of prenatal visits is roughly 50% lower for the Arizona sample, although the shape of the quantile-effect curve is extremely similar. The shape of the no-prenatal-care effect is also very similar to that of Washington, but the estimated panel effects are not significantly different from zero at any of the quantiles.

Hypothesis testing (Washington and Arizona):

- Minimum-distance testing framework based upon

$$\widehat{\gamma}^R = \underset{\gamma^R \in \Theta}{\arg\min} \left(\widehat{\gamma} - R\gamma^R\right)' \widehat{A}^{-1} \left(\widehat{\gamma} - R\gamma^R\right)$$

such that

$$\left(\widehat{\gamma} - R\widehat{\gamma}^R\right)' \widehat{A}^{-1} \left(\widehat{\gamma} - R\widehat{\gamma}^R\right) \underset{H_0}{\xrightarrow{d}} \chi^2_M$$

where $M$ is the number of restrictions (i.e., $M = rows\,(R) - columns\,(R)$).

- *Test of correlated random effects:* To determine whether a "pure" random effects specification (in which $c_m$ is uncorrelated with $x_m$) would be rejected for a given quantile $\tau$, the null hypothesis

$$H_0 : \boldsymbol{\lambda}_\tau^1 = \boldsymbol{\lambda}_\tau^2 = 0$$

is tested.

  - For each of the quantiles ($\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$) the null hypothesis is overwhelmingly rejected with a p-value extremely close to zero.

- *Test of the equality of the "effect vector" across quantiles:* This test considers whether there are *any* statistically significant differences in the $\beta_\tau$ estimates across two different quantiles.
  - The Table summarizes the results of this test applied to every pair-wise combination of quantiles from the set $(\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\})$
.

Table 2: Pairwise Tests of $\beta_T$ Equality Across Quantiles (p-values reported)

Panel Specification (Washington)

| | 10% | 25% | 50% | 75% |
|---|---|---|---|---|
| 25% | 0.000 | | | |
| 50% | 0.000 | 0.015 | | |
| 75% | 0.000 | 0.002 | 0.020 | |
| 90% | 0.000 | 0.000 | 0.000 | 0.012 |

Panel Specification (Arizona)

| | 10% | 25% | 50% | 75% |
|---|---|---|---|---|
| 25% | 0.000 | | | |
| 50% | 0.000 | 0.001 | | |
| 75% | 0.000 | 0.000 | 0.061 | |
| 90% | 0.000 | 0.000 | 0.231 | 0.859 |

- *Test of the equality of individual variables' effects across quantiles:* For a given variable (for example, marital status), this test checks whether the estimated effects at different quantiles are significantly different.

  – For the marriage indicator, for instance, the null hypothesis would be

$$H_0 : \beta^{married}_{\tau=0.10} = \beta^{married}_{\tau=0.25} = \beta^{married}_{\tau=0.50} = \beta^{married}_{\tau=0.75} = \beta^{married}_{\tau=0.90}.$$

Table 3: Testing Equality of Marginal Effects Across Quantiles

| | Washington | | Arizona | |
|---|---|---|---|---|
| | Cross Section | Panel Data | Cross Section | Panel Data |
| Second child | 0.121 | 0.057 | 0.000 | 0.061 |
| Male child | 0.000 | 0.000 | 0.000 | 0.000 |
| Age, $Age^2$ jointly | 0.010 | 0.246 | 0.000 | 0.450 |
| Education, $Education^2$ jointly | 0.012 | 0.358 | 0.001 | 0.946 |
| Married | 0.521 | 0.451 | 0.677 | 0.705 |
| No prenatal care | 0.013 | 0.005 | 0.359 | 0.867 |
| 2nd-trimester care | 0.573 | 0.095 | — | — |
| 3nd-trimester care | 0.109 | 0.610 | — | — |
| Smoke | 0.396 | 0.045 | 0.160 | 0.976 |
| Drink | 0.318 | 0.429 | 0.327 | 0.834 |
| # prenatal visits | 0.004 | 0.000 | 0.010 | 0.000 |
| Year of birth | 0.512 | 0.642 | 0.959 | 0.073 |

# Concluding remarks

- This paper has considered estimation of the effects of various prenatal-care variables and maternal characteristics upon quantiles of the (conditional) birthweight distribution.

- To deal with the unobserved heterogeneity of childbearing women, a panel dataset consisting of maternally-linked births was utilized.

- The estimated conditional quantile effects are analogous to the conditional expectation effects that arise from the correlated random-effects model of Chamberlain (1982, 1984).

- Since the quantile-regression techniques (and testing methodology) are straightforward to apply and the estimated effects have a rather simple interpretation, the approach of this paper should be useful for other researchers seeking to estimate "causal" quantile effects through the use of panel data