PRELIMINARY DRAFT – PLEASE DO NOT QUOTE

The Effect of Birth Order on Early Educational Attainment

Jennifer Gerner
Department of Policy Analysis and Management
Cornell University
jlg19@cornell.edu

Dean R. Lillard
Department of Policy Analysis and Management
Cornell University
DRL3@cornell.edu

Prepared for presentation at the 2006 annual meetings of the
Population Association of America
Los Angeles, CA March 2006

Acknowledgements
We thank Eamon Molloy for expert research assistance.

Abstract
We explore whether and how educational performance varies with birth order. Starting from the simple associations between birth order and educational performance, we show that these differences persist when one controls for family size and other demographic characteristics and that the birth order differences persist over time. We then argue that birth order matters for other important decisions parents make that are important inputs into the educational production process. In particular we argue that birth order affects a mother's decision to participate in the labor force, the decision about whether to enroll her child in pre-kindergarten, and the decision about the age she decides to enter her child in formal schooling. To account for the role these decisions play on how birth order affects ultimate educational performance, we use cross-state and temporal variation in compulsory schooling laws and state labor market conditions. Because our data, the Children of the NLSY79, cover a long time period (1986 through 2002) we observe a large number of children within families who were allowed to enter school at different ages. These differences in the age of first permitted entry mean that parents face greater incentives to enter a youngest child into school at an earlier age than they did an older sibling. We estimate a model of six simultaneous equations to account for the way in which birth order affects each. The results suggest that, even controlling for the association between birth order and other behaviors that influence educational attainment, birth order effects persist in statistically and economically meaningful ways. Children of higher order birth do less well in on achievement tests. We also estimate value-added models of gains in achievement. We find that birth order penalties disappear in mathematics but persist in tests of reading and picture-vocabulary association. The results suggest what commensense tells us – that actors in the educational production process likely reallocate resources in favor of children who are underperforming in school.

I. Introduction

Although social science researchers have long been interested in the study of how education gets produced, the passage of the "No Child Left Behind" Act heightened awareness of and general interest in trying to understand how educational performance varies with individual and family background characteristics. Because the Act (essentially) requires all children to meet one standard, researchers have redoubled efforts to understand the role played by factors that may influence a child's initial readiness or ability to learn. This focus is appropriate because ultimate human capital production builds on foundations laid early in life. Consequently, researchers are paying special attention to factors related to early educational production. These include factors that policy can easily influence such as the age by which a child must be enrolled in school and whether or not kindergarten or pre-kindergarten programs are universally offered (required). Researchers are also reviving a long running interest in the role that family size, birth order, and parental choices play on the early educational performance of children.

In this paper, we focus on this second set of factors but make use of cross-state and temporal variation in state policies governing when a child must attend school. Our primary question is whether a child's early educational varies with the order he was born into his family. While this question is part of a well-ploughed field, researchers are bringing to it new tools and new data yielding new insights. In this paper we bring to bear new data and a reframing of the analytical approach that advances our understanding of the role that birth order might play in determining early educational performance. We model not only the variation in educational performance but also account explicitly for how birth order is correlated with parental decisions about other inputs to the production of education. We focus on the decision by parents to enroll their child in pre-kindergarten, the decision by mothers to participate in the labor force, and the choice of parents about when to enroll their child in formal schooling. Because our analysis allows these decisions to be made jointly, we provide evidence from a richer investigation. We also contribute because we use multiple sources of variation to identify these decisions. Consequently, we provide a number of tests of the assumptions on which the investigation rests.

In what follows we first describe the raw association between birth order and performance on standardized tests in mathematics and reading taken over six early years of a child's formal schooling. We then show how the association between birth order and educational performance changes when one controls for family size, individual and family background characteristics. Our basic findings are that, relative to first born children, children born later perform worse on all subjects on which they are tested. These differences persist when one controls for the number of children in a family and for basic demographic characteristics. While the penalty of birth order persists over time in most subjects, it is wiped out relatively quickly in mathematics. We confirm this finding by estimating multivariate regression models of the change in scores of a given child across tests separated by between two and four years. The results of those models are consistent with the hypothesis that performance differences associated with birth order narrow over time. This narrowing of differences is consistent with a dynamic model of investment in human capital in which actors (parents, teachers, children) reallocate resources to compensate for deficiencies in learning.

II. Background

To start, we plot the raw association between birth order and several measures of educational attainment. While we describe all the data in detail below, a picture of the raw mean differences in educational performance of children of different birth order helps set the stage. In Figures 1-3 we plot data on the educational performance of children of different birth order on three sets of tests administered over six years (roughly two years apart) starting about the time the children first entered school. The outcome data consist of the percentile rank of the score a child achieved on tests of mathematics, reading recognition, reading comprehension, and a picture-word association test that were administered as part of the Children of the National Longitudinal Surveys of Youth (CNLSY).

Figures 1-3 show what is commonly believed - that later born children do not perform as well in school as their older siblings. Figure 1 shows that the test scores of a first born child place him roughly 10 to 20 percentile rank points higher than the percentile rank of a child born fifth or higher in the birth order. Figure 2 and Figure 3 show that these raw differences persist over time even as children gain experience in school. Test scores of first born and fifth or later born children on the second and third administrations of the standardized tests (roughly two and four years later) result in percentile rankings that continue to differ by 10 to 20 percentile points.

The patterns shown in Figures 1-3 have figured in a wide swath of social science literatures that includes demography, sociology, psychology, and economics. Conceptually many birth order studies appeal to the idea, attributed to Becker (1960), that parents make decisions about both the number of children they will have as well as the resources (time, effort, money) they will invest in each child. The theory advanced by Becker and extended by others (Becker and Lewis, 1973; Willis, 1973) sought to explain secular changes in family size. However, embedded in the model is not only the idea that parents trade the average quality of their children against the number they have but also the idea that parents implicitly trade the quality of a given child against the quality of his sibling.

This notion is more transparent when framed in the context of an educational production function in which parents (and teachers) allocate resources to and across children to maximize the human capital each child acquires (Hanushek, 1979). In that framework parents explicitly choose to allocate time and effort across children. Because available parental time is physically constrained, when parents choose to bear multiple children, the children on the ends of the birth order (first and last born) get more parental resources than do children in the middle of the birth order.

In numerous empirical studies, demographers, sociologists, psychologists, and economists have investigated the association between birth order and educational achievement. Studies differ in the set of factors they hold constant but most studies include family size and/or the interval between siblings. Although some studies have found that ability and achievement appears to decline with birth order, even holding constant family size (Steelman, 1985; Cicirelli, 1978), others find that this is an artifact which disappears when a range of other relevant variables are held constant (Hauser and Sewell, 1985). In an economic study framed by the educational production function approach, Hanushek (1992) finds that although it is always better to be in a smaller family, there is no particular advantage associated with birth order. Instead, he argues, first born children outperform second born children because first born children live families that are, on average, smaller than families of second born children. After accounting for the negative relationship between family size and educational performance, Hanushek (1992) finds no evidence of birth order effects on educational performance.

Black, Devereux, and Salvanes (2005) try to disentangle the effects of birth order from family size using Norwegian administrative data. They examine whether children born in a particular birth order attain more schooling. To separate effects of birth order from the effects of family size they compare children who are twins with single born children of the same birth order. Under the assumption that the arrival of twins represents an exogenous shock to family size, the comparison of educational attainment of twins to non twins in the same birth order separately identifies birth order effects. They find a negative correlation between family size and years of schooling, but this seems to be driven by birth order effects.

The empirical birth order literature is now moving toward an explicit recognition that any exogenous allocation of resources to first born children is transitory and may be undone if parents reallocate resources in favor of later born children. This recognition flows from the fact that, nine to twelve months after the birth of each child, parents again begin to *choose* whether or not to have another child and, after the birth of second and every other higher order born child, parents *choose* whether to reallocate resources across children of different parity. The fact that parents are able to choose means that empirical studies of birth order effects must account for the choices parents make that are correlated both with birth order and with a child's educational performance. Among others, these decisions include: whether or not to enroll a child in pre-kindergarten programs, whether or not (when and how much) a mother will work, and the age at which a child begins his formal schooling. While studies exist that examine how these decisions are correlated with educational performance of children, no study ties together these decisions in the context of how they vary with and are determined by birth order.

Empirical studies on the effects of pre-kindergarten participation generally suggest that early participation in high quality programs can generate positive improvements in later school performance and other child outcomes. Evaluations of Head Start (Currie and Duncan, 2000), the Perry Pre-kindergarten Study (Schweinhart, Barnes, and Weikart, 1993), and other similar interventions (Cunha, Heckman, Lochner, and Masterov, 2005) show positive effects at least for some groups of children. More recently, Gormley, Gayer, Phillips and Dawson (2004) evaluate gains in school performance of children who participated in a broadly available pre-kindergarten program in Oklahoma. They find participation raises short term performance. While the results of these studies are provocative, it is difficult to draw general inferences about what might happen if such programs were mandated because the evidence is based on program data where parents were not required to enroll their children. Because the families in the above studies choose whether or not to participate the usual issues about self-selection must be confronted.[1]

Empirical studies yield mixed evidence about whether school performance is related to the age a child first enters school. These studies suffer in part because there is no strong consensus about what the theoretical effects will be if a child enters formal

[1] Recently policy makers in several states have suggested that states should provide universal state-funded pre-kindergarten programs.

schooling at a younger or older age. On the one hand, school performance should be higher among children who enter school at younger ages if early school entry exposes them to a richer learning environment than they would get at home. On the other hand, school performance is likely to suffer among early school entrants who are not emotionally "ready" for school (Kagan, 1990).

Available empirical evidence suggests that both associations are present but there is scant evidence about the direction of the causality or about the net association between early school entry and subsequent educational performance. For example, Crosser, (1991) and Stipek and Byler (2001) find that children whose parents enroll them earlier perform worse. By contrast, DeMeis and Stearns (1992) and Graue and DiPerna (2000) find that performance either does not differ across children enrolled at different ages or is even slightly higher when parents enroll their children at a younger age.

As in the literature on pre-kindergarten attendance, this literature largely fails to account for the endogenous choices of families about when to enroll a child in formal schooling. Given that parents observe much more about a child than do analysts, the self-selection issues are important to consider before any firm conclusions can be drawn. Recently Datar (2003) has accounted for this endogeneity by using school district policies as an instrument for age at enrollment. She finds that a one year delay in starting school increases subsequent school performance.

The question of how birth order affects educational performance is further complicated by questions of whether birth order affects a woman's decision to participate in the labor force. This decision is clearly related to the decision to enroll a child in pre-kindergarten and the decision about when to enroll a child in formal schooling. It is likely that these three decisions are taken (more or less) jointly.

Consider the simple correlation between birth order, the probability a child participated in pre-kindergarten and the probability his mother worked in the year he entered first grade. In Figure 4 we plot the fraction of children of each birth order who were enrolled in some kind of pre-kindergarten program (Head Start, nursery school or other pre-kindergarten program) and the fraction of children of each birth order whose mothers worked in the year they entered first grade.

In both cases it is apparent that birth order is also correlated with the probability a child experienced pre-kindergarten and with the probability that a mother worked (or how much she worked). The probability of pre-kindergarten and a working mother for first born children is more than 20 percentage points higher than it is for children born fifth or later in the birth order. While these bivariate associations are unsurprising - mothers of larger families tend to specialize in home production – they suggest that one should consider whether and how birth order affects the joint decision of a mother to enter, stay in, or exit the labor force and enroll her child in (pre) school at a given age.

While no studies directly confront how all three decisions are related to birth order, studies do show that entry into school, public school enrollment, and labor force participation of women are related. For example, Datar (2006) has looked at the joint determination of child entry into school and mother's labor force participation. She is primarily concerned with how changes in school entry policies affect a mother's labor force activity and child care decisions. She shows that kindergarten, indeed, provides a source of child care. When policies restrict entry to older ages, women participate less and spend more money on (private) child care. Gelbach (2002) examines whether single mothers are more likely to enroll their youngest child in public school as soon as the child is age eligible to be enrolled in public school. He finds that single mothers are more likely to participate in the labor force when their youngest child is age eligible to be enrolled in school. Gelbach (2002) estimates that, when parents are permitted to enroll a child in public school one year earlier, single mothers increase their labor supply by between 6 and 24 percent.

III. Conceptual Aspects of Birth Order

We investigate whether a child's birth order affects his school performance, controlling explicitly for related parental decisions that may also be affected by birth order. We adopt the education production function framework to estimate performance on standardized tests in mathematics and reading. We follow Hanushek (1992) in assuming that an individual's performance on an educational achievment test (A_{it}) is a function of his past achievement (A_{it-1}), family inputs (F_{it}), school inputs (S_{it}), and other exogenous (possibly time-varying) factors (X_{it}). Although birth order (B_i) is exogenously assigned to an individual, we list it separately because it is our covariate of interest. We specify individual i's performance on a particular test administered at time t as:

(1) $A_{it} = \beta_0 + \beta_1 A_{it-1} + \beta_2 F_{it} + \beta_3 S_{it} + \beta_4 X_{it} + \beta_5 B_t + \varepsilon_{it}$

where ε_{it} is a person specific stochastic shock to performance that varies over time.

To formalize the above observations we modify (1) to explicitly build in a role for endogenously chosen determinants of past achievement (A_{it-1}) and family inputs (F_{it}). We model past achievement as a function of whether a child participated in pre-kindergarten (PS_{it-1}), the age he entered formal schooling (Age_{it-1}), and the time his mother spent at him with him (L_{it-1}). In our analysis past achievement represents human capital production that occurred before grade 1. It will therefore include production that occurred both in the home and in other settings outside the family. The inclusion of both pre-kindergarten attendance and past labor force participation of mothers is meant to proxy for the production of past educational attainment. Current achievement depends on the time a mother currently spends working in the labor market (L_{it}). Finally, we modify (1) to formalize the conjecture that birth order may also enter the achievement function indirectly because it influences whether parents enroll a child in pre-kindergarten and the decision of the mother to participate in the labor force. The modified achievement equation is given by:

(2)

$$A_{it} = \delta_0 + \delta_1 A_{it-1}(PS_{it-1}(B_i), Age_{it-1}(B_i), L_{it-1}(B_i)) + \delta_2 F_{it}(L_{it}(B_i)) + \delta_3 S_{it} + \delta_4 X_{it} + \delta_5 B_t + \varepsilon_{it}$$

We have modeled achievement in a way that allows birth order to directly and indirectly influence achievement. To estimate the direct effect of birth order (i.e. to estimate the elements of δ_5) we must account for the effect of birth order on the other determinants of achievement. We estimate (3) with both OLS and with instrumental variable methods (including birth order as a determinant of each outcome). We estimate these decisions in a simultaneous equations framework to allow the errors in each outcome to be correlated.

Instrumental Variables

To estimate the direct association between birth order and educational performance we need to estimate the indirect association between birth order and performance through other behavior of parents that potentially influence test performance. Since parents choose whether to send a child to pre-kindergarten and kindergarten, the age to enroll the child in formal schooling, and the mother's labor force participation we use the method of Instrumental Variables in a system of simultaneous equations that allow the error terms in each equation to be correlated. The method of Instrumental Variables (IV) is a technique described in both the economics and sociology literature and provides a method of accounting for unobserved heterogeneity (Meyer, 1995; Winship and Morgan, 1999). This method is more intuitively understood if one thinks of it as consisting of two stages. In the first stage, one estimates each of the above parental decisions that are determined partly by a child's birth order. The method of instrumental variables requires a set of factors (Z_{it}) that determine the outcome of interest but that are uncorrelated with the child's school performance.

Using the mother's labor force participation as an example, the application of instrumental variables requires us to predict her labor force decision using birth order (and other covariates) plus the instrumental variables. That is, we would estimate a first-stage equation that generally takes the form:

(3) $L_{it} = \alpha_0 + \alpha_1 B_i + \alpha_2 Z_{it} + \alpha_3 X_{it} + \eta_{it}$

The effect of a mother's labor force participation on the educational outcome is identified by variation in the instrumental variables (Z_{it}) that is independent of those components of the error term (η_{it}) that influence a child's educational attainment. Our instrument set, described in more detail below, includes the ages set by state statute that govern the age a parent may first enroll her child in public school, the age by which she must enroll the child in school, a set of factors describing the conditions under which states exempt parents from the compulsory school age, and conditions of labor markets in the state.

Challenges to the exclusion restrictions

In principle, it is easy to defend the exclusion of these variables from the equation describing a child's educational attainment because, other than their effects on actions parents take, it seems implausible to argue that these variables should directly determine a child's school performance. In practice, however, there may be incidental correlation between state labor market conditions, compulsory schooling laws, and unobserved determinants of school performance (*i.e.*, the error term ε_{it} in equation (2)). If the explanatory power of the instruments is weak, even seemingly small incidental correlation can cause severe inconsistency in the IV estimator (Wooldridge 2002, pp. 101-102). Incidental correlation might arise, for example, if state compulsory schooling ages are set in conjunction with the allocation of public resources to education. Since the quality of a state's public education system probably does influence whether or not parents choose to move to that state, a correlation between the state's compulsory schooling laws and the funding of public education threatens the validity of the exclusion restriction.

Available evidence suggests that this potential threat to validity is unlikely. Numerous studies that use compulsory schooling laws to predict educational attainment find a strong correlation between schooling and complusory schooling laws in both the US (Adams, 2002; Lleras-Muney, 2005) and internationally (Arendt 2004; Chou et al., 2004; Spasojevic, 2003). These studies find that the compulsory schooling laws predict levels of educational attainment well but do not directly investigate through what channel the effect works. A few studies use compulsory schooling laws to directly estimate their effect on the age a child enters school using data from the US (Gerner and Lillard, 2005) and other countries (Fertig and Kluve, 2005; Frederiksson and Öckert, 2004; Jürges and Schneider, 2006). These studies uniformly find that compulsory schooling laws strongly predict the age children enter formal schooling.

As a robustness check on our identification of parental decisions and to try to overcome this possible challenge we further refine our estimation strategy to account for unobserved time invariant factors that may be correlated with state compulsory schooling laws or state labor market conditions and that might also influence a child's school performance.

Introducing subscripts f to denote a given family and subscript s to deonte a particular state, we allow the error in (2) to be of the following form:

(4) $\varepsilon_{ifst} = \nu_i + \varphi_f + \mu_s + \omega_t + \zeta_{ifst}$

where ν_i, φ_f, and μ_s are unobserved factors that are constant over time for individuals, families, and states respectively. Analysts often interpret the person-specific fixed component (ν_i) to represent person specific ability but it may more generally represent any unobserved time-invariant factor specific to an individual that influences his achievement. The component φ_f is often interpeted to represent factors such as family-specific attitudes towards education and other similar time-invariant factors that affect a child's school performance. The component μ_s captures time-invariant factors specific to each state (such as the willingness of voters to fund public education). The component ω_t captures cohort effects. ζ_{ifst} is the classical error term that is randomly distributed across persons, families, states, and time.

We estimate models with various combinations of restrictions on the errors in (4).[2] We estimate a model with only state fixed effects, a model with only family fixed effects, and a model with both family and state fixed effects. That is we estimate the modified empirical counterpart of (2) as:

(2′) $A_{ifst} = \delta_0 + \delta_1$ **Predicted PreSchool**$_{ifst} + \delta_2$**Predicted LFP of Mother**$_{ifst}$

$+ \delta_3$ **Predicted First Grade Age**$_{ifst} + \delta_4$ **Age at Assessment**$_{ifst}$

$+ \delta_5$ **Birth Order**$_i + \delta_6$ ***Individual Background***$_{ifst} + \delta_7$ ***Family inputs***$_{ifst}$

$+ \delta_8$ **Year began school**$_{ifs} + \delta_9$ ***Family Fixed Effects*** $+ \delta_{10}$ ***State Fixed Effects*** $+ \nu_i + \zeta_{ifst}$

We estimate unrestricted and restricted versions of (2′). These correspond to models in which we test whether or not attainment is influenced by time-invariant effects of unobserved factors of either a state or family. Formally, we estimate models where we alternatively restrict $\delta_9 = 0$, $\delta_{10} = 0$, and both $\delta_9 = 0 = \delta_{10}$.

To sweep out the influence of the person-specific fixed effect (ν_i) we also estimate models of the change in test scores from one administration to the next. This model estimates the gain in percentile rank from one administration of the test to the next as a function of time varying factors related to school performance. We include in that model birth order and other time invariant factors such as race because we posit they are correlated with time-varying factors that influence not only the level but also the trajectory of human capital production. The coefficients on birth order constitute a test of the hypothesis about whether factors associated with birth order affect how much children learn.

[2]We always control for the year a child entered first grade to allow for secular trends in school performance.

The value added models will take the general form:

(5) $A_{ifst+k} - A_{ifst} = \varphi_0 + \varphi_1$ (**Age at Assessment**$_{ifst+k}$ **- Age at Assessment**$_{ifst}$)

$+ \varphi_2$ **Birth Order**$_i$ $+ \varphi_3$ ***Individual Background***$_{ifst}$

$+ \varphi_4$ (***Family inputs***$_{ifst+k}$ **-** ***Family inputs***$_{ifst}$) $+ \zeta_{ifst}$

Note that this value-added model does not solve the problem of all bias associated with unobserved factors. This model only removes bias associated with those unobserved time-invariant factors that influence *the level but not the change* in educational achievement between two time periods. Any time-invariant factor that influences the change in educational achievement remains a source of potential bias. The value-added model has the additional disadvantage that it requires much from the data. With these caveats, the value-added model provides a different test of the proposition that birth order matters – one that takes care of some but not all potential bias due to unobserved heterogeneity. Note also that, conceptually the family inputs, time devoted by parents, resources available in the home, family size, composition of the household, all vary over time and may influence the change in percentile rank from one administration of the test to the next. We are able to include some time varying family background factors in the model we estimate.

Sources of identification

To identify the above models, we use variation in compulsory schooling policies and state labor market conditions across states and over time in the same states. In addition, we are able to use variation in policies and conditions that arises because families moved across state lines in relevant years. For the compulsory schooling policies, this means that families moved across state lines in a year such that an older child faced a different compulsory schooling age than one of his younger siblings did. All cross-state moves imply that parents and mothers faced different labor market conditions. For the sample of children who did not move across state lines, we take advantage of two sources of identification for the preschool enrollment and age entered first grade equations. The first source of variation is commonly used – variation that arises because a child is born before or after the month that a state used to decide who would or would not be permitted (or required) to enter school that year. The second source is much less commonly used – changes within a state over time in the compulsory schooling laws. Because our sample period stretches from 1977 to 2001 we are able to exploit changes in state compulsory schooling laws over time within a state. We also exploit these sources of variation across siblings in the same family to the extent allowed by our sample.

Identifying a causal effect of family size

One of the continuing debates in the birth order literature is whether birth order affects a child's educational attainment independently of family size. As noted above, the evidence on this question remains mixed. Hanushek (1992) finds no independent effect of birth order but all variation in family size that he observes in his data is endogenous. Black et al (2005) assume that the arrival of a twin represents and unanticipated increase in family size. Using administrative data covering the population of Norway, they do find evidence of an independent effect of birth order on educational attainment.

We adapt the empirical strategy of Black et al (2005) in our analysis. In particular, we make use of the birth of 115 sets of twins and 2 sets of triplets in the families in our sample. We use the arrival of the second twin and the last two of the triplets as exogenous shocks to family size. Under the assumption that these multiple births were unanticipated, their arrival constitutes a shock to family size that affected all siblings alive before these multiple births and the twins and triplets themselves.

IV. Data

We use data from the National Longitudinal Survey of Youth 1979 (NLSY79) and data from the Children of the NLSY79 surveys (CNLSY79). The NLSY79 cohort originally included 12,868 young men and women who were 14 to 21 years of age as of December 31, 1978. Annual interviews were conducted for these men and women beginning in 1979, shifting to biennial interviews beginning in 1994. Extensive information about education, employment, training and family experiences has been collected from this sample. Starting in 1986 data have been collected from the children of female respondents of the NLSY79. To be interviewed, young children (ages 0 to 14) must live at least part time with their mothers. The CNLSY79 surveys include data from biennial assessments of cognitive ability, the home environment, temperament, behavioral problems, self competence, and other information. Because all the children of the women of NLSY79 are interviewed, the data include a number of siblings from the same family.

Measures of school performance

The measures of school performance we use are scores children received on the Peabody Individual Achievement Tests (PIAT) in mathematics, reading recognition and reading comprehension and the Peabody Picture Vocabulary Test (PPVT). The PIAT and the PPVT, administered three times to each child between the age of 5 and 14, are brief assessments of academic achievement. These widely used tests have high test-retest reliability and concurrent validity. The PIAT-Math assessment consists of 84 multiple-choice items assessing recognition of numerals and progressing to measuring advanced concepts in geometry and trigonometry. The PIAT-Reading Recognition measures word recognition and pronunciation. Skills include matching letters, naming names, and reading single words aloud. The PIAT Reading Comprehension assessment measures a child's ability to derive meaning from sentences that are read silently. This assessment is first administered after children have attained some competency in reading recognition. The PPVT provides an estimate of verbal ability or scholastic aptitude. It is administered to NLSY79 children between the ages of 3 and 18. Scores on each test are reported in terms of raw scores, scores that have been standardized for age and grade in school, and the percentile rank of each child's score. In all of our analyses, we use the percentile rank of each child's score.

Table 1 shows the means and standard deviation of the percentile ranking of the test score on each test, each time it was administered to the CNLSY79 sample. It also shows the mean and standard deviation of the childrens' age (in months) when they took the test. The average child first took the mathematics and reading recognition tests when he was 78 months old (about 6.5 years old). The PPVT was first administered when the average child was about one year younger (66 months old). As noted before, the Reading Comprehension test was administered later, after some reading skills had been developed. It was first administered when the average child taking the test was about 92 months old (7.7 years old). With the exception of the first and second administration of the PPVT, each successive administration of these tests took place about two years later. The PPVT was administered when the children were about 116 months old (about 9.7 years old) or about four years after it was first administered.

Our main covariate of interest is the birth order a child occupies among his siblings. To simplify the analysis, facilitate the interpretation of the results, and to avoid imposing arbitrary functional form assumptions, we create six variables that indicate a child's order in the birth of the children in a family. In all cases the reference group is first born children. The six variables identify children who were born as a singleton (no other sibling), born 2nd, 3rd, 4th, and 5th or higher in the birth order and those who are the youngest in their family (last born).

Age entered first grade

To assign each child the age he or she first entered first grade, we combine information from the NLSY79 household record data with information on each child's sex and date of birth. In particular, we first link children in our CNLSY79 sample to information on each member of the household reported annually by the NLSY79 respondents. Each time an NLSY79 respondent is surveyed, he or she reports the age, sex, and highest grade that each household member has completed as of the date of the interview. Because we know each child's sex, year, and month of birth, we can match each respondent in the CNLSY79 to the corresponding information on that child as reported by the mother in the NLSY79 household composition files. Using the date of each interview, this match allows us to observe the grade completed by each child in the CNLSY79 annually from 1979-1994 and biennially thereafter. We then infer the school year in which a child entered first grade by noting the year the mother reported her child had completed either kindergarten or first grade.

In the survey years from 1994 onwards, we inferred the date the child entered first grade as follows: If the child completed kindergarten in a survey year (i.e. in 1994, 1996, 1998, 2000, or 2002) we assumed the child entered first grade in September of that calendar year. If the child completed first grade in the NLSY79 survey year and he was not enrolled in a previous survey year, we assumed he entered first grade in the fall preceding the calendar year of the NLSY79 survey.

Note that this algorithm introduces systematic measurement error into the assigned age of entry in first grade for children entering school in 1995 through 2002 because we do not observe the grade a child completed in the year his mother was not interviewed. Consider the nature of the error in the context of a particular survey year – 1996 – that followed a year in which no survey was administered. Suppose we observe a child to have completed first grade in the spring of 1996. If the most recent past survey (1994) does not show that child completed first grade in 1994 then we are faced with three possible assignments for the date he entered first grade. The child might have entered first grade in the fall of 1994, failed the grade, and then repeated it to complete first grade in the spring of 1996. In this case we would assign him a (first grade) entry date of September 1995, one year after the date he actually entered first grade. The child might have also entered directly into first grade in the fall of 1995. In this second case, our algorithm assigns him the correct date of entry into first grade. Although it is difficult to find national data on the fraction of first grade students who repeat a grade, Shepard and Smith (1989) report that the fraction of first graders that repeated the grade in the 1985-86 school year in 12 states and the District of Columbia range from 1.6 to 20 percent. The population weighted average first grade repetition rate in this convenience sample was 11.6 percent. Using a national sample of more recent data, Zill and West (2001) find that the kindergarten retention rate of first and second grade students was 6 percent in 1993 and 5 percent in 1995. These data suggest that, we will have misassigned the age a child entered first grade for between 5 to 12 percent of the sample who entered school after 1994.[3]

Pre-kindergarten attendance

Our measure of pre-kindergarten attendance is an indicator of whether the child ever attended some type of schooling before entering formal schooling. These data are drawn from questions, asked of everyone in the NLSY79 surveys, that indicate whether a child "is currently attending..." and whether the child "ever attended pre-kindergarten/nursery school."

Mother's labor force participation

We construct measures of the mother's labor force participation using employment data from the NLSY79 surveys. We use all available employment data, including retrospectively reported data on the dates a women started or stopped working. Using those data we construct annual indicators of whether a woman worked in each calendar year. We then construct two measures of a woman's labor force participation. The first measure is a count of the number of years a child's mother worked between the year he was born and the year he turned 5. The second measure is an indicator of whether or not a mother moved into the labor force (from not working) in the year the child entered first grade.

Exogenous Shocks to Family size

[3] We corrected our assignment with available data from the CNLSY79 School Supplement Survey on whether a child ever repeated or skipped kindergarten or first grade but these data only cover 3620 children and were collected only in the 1993-94 and 1994-95 school years.

We code a variable that indicates the number by which family size exogenously increased using data on the birth of twins and triplets. If twins were born in a given year then family size for all children alive at that time (including the twins) is exogenously higher by 1 person. If triplets are born, family size is higher by two persons. The family size shock variable remains constant at the assigned value until the next child is born (at which time it is reset to be zero if only one child is born).

Other covariates

We include a parsimonious set of control variables. These include the number of siblings the child has, the child's age (in months) when he took the standardized test, his race (black, other), the number of adults present in the household, the highest grade his mother completed, her marital status, the highest grade his maternal grandmother completed, and his mother's Rotter index score on the locus of self-control questions in the NLSY79 survey. The Rotter index of locus of control is based on four questions designed to measure the extent to which people believe they have control over their own lives, as opposed to having lives which are dictated by external forces such as chance or fate.[4] We include it as a potentially important control variable somewhat related to the idea of non-cognitive skills (Heckman and Rubenstein 2001), in this case of the mother, that may influence her childrens' school performance. Table 2 reports the means and standard deviation of the above dependent and independent variables. Sample sizes vary across the different administrations of each subject test. We do not report the mean and standard deviation for each of the samples but those results are available on request.

Data used as instruments

[4] HIgher Rotter index scores indicate less feeling of control over one's life.

Our set of instruments includes data on compulsory schooling laws that govern the age parents may first enroll and finally must enroll a child in school and variables measuring labor market conditions. All of these variables are defined at the state level.

The data on state compulsory laws, the Compulsory Schooling Law (CoSLAW) database, were compiled by the authors. These data cover the complete history of compulsory education laws for each state. This compilation of these data was facilitated by the presence at Cornell University of one of the most complete collections of state statutes in the United States. We have coded all changes in laws that specify the age a state first permits parents to enroll a child in public school, the age by which parents must finally enroll a child in (public or private) school, the conditions under which exceptions are granted to these ages, and penalties faced by children, parents, and school officials if a child fails to attend school. In addition, these data include information about how these laws are enforced and financed. While other compilations of such data are available, those compilations are typically cross-sectional snapshots taken every five years or so. What makes our data unique is that we have coded the exact date specific changes to various provisions of the laws took effect. Thus we can exploit the full variation available in these data. For every state, we standardized compulsory school entry ages so they indicate how old a child had to be on September 1st of a given year.

State labor market variables

Data on state labor market conditions are drawn from the Bureau of Economic Analysis series SA05 "Personal Income by Major Source and Earnings by Industry" available online. This series includes annual total earnings in the industries in each state (defined by one-digit Standard Industrial Classifcation code). We use the data to construct the log of per capita earnings in agriculture, mining, construction, manufacturing, wholesale trade, retail trade, finance (including insurance and real estate industries), services, civilian employment in federal government jobs, military, and state and local government. Data on youth and adult unemployment rates were drawn from the US Bureau of Labor Statistics annual publication *Geographic Profile of Employment and Unemployment*.

Matching of instruments to CNLSY79 respondents

To match these instruments to CNLSY79 respondents we use the restricted access data files of the NLSY79 that includes state geocode identifiers of the given year. We use state geocodes to identify the compulsory schooling laws faced by each child in each year. We use variation in the laws to identify whether or not a mother enrolled her child in pre-kindergarten schooling, in kindergarten, and the age (year) he was when he was enrolled in first grade for the first time.[5]

We rely on different sources of variation in the instrumental variables to identify our models. We use variation in the compulsory schooling ages across states, over time, across siblings in the same family for families that did or did not move across state lines in the relevant years. To get a sense of the sample that is subject to different levels of those variables, we use the sample of children with at least one other sibling and for whom we have a valid test scores for both children from the first administration of the PIAT mathematics test. Table 3 shows the mean and standard deviation of each of the compulsory schooling age for the sample of children with siblings whose families moved across state lines. The lower half of Table 3 shows how many children faced a different

[5]About one third of our sample lived in states that did not specify an age parents were permitted to first enroll their child in public school. In such cases we set the permitted age of entry equal to the compulsory age of entry. We code an indicator variable to equal one if age was imputed and includ it in all regressions.

compulsory schooling law than his siblings. The data here show that about 1200 children faced a different compulsory schooling age than at least one of their siblings, evenly divided by whether they faced an older or younger age than their siblings. These numbers suggest it will be possible to estimate models with family fixed effects.

V. Results

The CNLSY data, the raw differences in percentile score rankings mostly persist even when one holds constant individual and family demographic characteristics, including the number of siblings a child has. Further, the estimated direct association between birth order and educational performance is generally higher when one accounts for the indirect association that operates because birth includes other related parental decisions. Tables 4-7 show how the percentile score ranking changes when one successively controls for number of siblings, for a set of demographic characteristics, and for endogenous parental behaviors. In the first column associated with each test administration (first, second, and third), we adjust the raw percentile score only for the number of siblings the youth has and the age at which he took the standardized test. In the second column we additionally control for individual and family background characteristics. In the third column we control for other endogenous parental behaviors that are likely affected by birth order.

The results in Tables 4-7 show that, even with the adjustments embedded in the different models, statistically significant differences remain in the percentile scores in each subject area and across all three administrations of the tests. In the most stringent empirical models, the percentile rank of scores of first born (non singleton) and fifth or later born children differ on average by between 3 to 10 percentile rank points.[6]

Relative to the raw differences shown in Figures 1-3 and Table 1, adding controls for family size reduces the penalty associated with being born higher in the birth order. The penalty falls further when one accounts for demographic differences across youth. By adding family background characteristics, the score gap between first born and higher born youth falls by between .5 and 9.3 percentile points (a reduction that constitutes between 10 and 200 percent of the base score gap). The difference between the estimates in column 2 and column 3 show the bias in estimated birth order associations from not accounting for the correlation between birth order and other endogenous parental choices. In some cases there appears to be little bias. In other cases (e.g. the first mathematics test and the PPVT), the bias is considerable.

Despite the reduction in the birth order score gap, statistically significant birth order score differences persist even with controls for family background and sibship size. After adding controls for family background, youth born second or later in the order score less on all four tests in all subjects. With only two exceptions, the penalty of being higher in the birth order remains statistically significant. While the penalty of being higher in the birth order remains surprisingly constant (or even grows) over successive administrations of the test, it appears that children of higher birth order "catch up" to their first born peers in mathematics.

Tables 8-10 show the results from the value-added or person fixed-effects models that correspond to equation (5). Table 8 shows the relationship between birth order and

[6] Appendix figures 1.1-1.3, 2.1-2.3, 3.1-3.3, and 4.1-4.3 show graphically how the penalty associated with birth order changes across these models.

the change in percentile rank on the first and second administration of each test. Table 9 shows the relationship between birth order and the change in percentile rank on the first and third administration of each test. Table 10 shows the relationship between birth order and the change in percentile rank on the second and third administration of each test. Note that the samples available to estimate the models are limited to children with valid scores on the relevant tests.

The results in Table 8 confirm the pattern observed in the models of the percentile rank on each test shown in Tables 4-7. Children of higher order birth narrow the gap in their percentile rank over time on the mathematics and on the PPVT tests. Children of higher order birth (relative to first borns) gain close the gap so that, as seen in Table 4, by the time the third test is administered, there is no significant difference in the percentile rank of the mathematics test score for children of different birth parity. Similarly, children of higher birth order gain more between the time the first test was administered (for most children the first PPVT was administered before they entered first grade) and the time the second PPVT was administered. Despite these gains, children of higher order birth occupy a lower percentile rank on the PPVT score distribution than do their first born siblings. Table 9 shows a similar pattern for mathematics test scores when one compares the gain in percentile rank between the first and third administration of each test. On all other tests the pattern suggests that children of higher birth parities close the gap in percentile rank but none of these estimated gains are significant at conventional thresholds of statistical significance.

A comparison of the reults in Table 8 and Table 10 suggest that the gains in mathematics performance occur early in a child's educational trajectory. One must interpret the results cautiously because the samples differ but the results are consistent with the proposition that parents and/or teachers differentially allocate attention (time and/or resources) towards children of higher birth order. The pattern suggests that actors in the educational production process adjust to overcome deficiencies associated with higher birth order that affect learning of mathematics. That the same pattern is not observed for the other tests remains to be explored.

VII. Discussion and conclusions

The question of birth order has long fascinated social scientists, in part because one is tempted to treat it as a random assignment. A child does not choose the order (or family for that matter) into which he is born. In this paper we note that, while birth order is exogenously assigned, parents and teachers can adjust their behavior to compensate for any differences in resources that a simple birth order assignment implies. Our analysis shows that, when deciding whether to take actions that potentially affect the educational attainment of their child, parents choose differently for children of different birth order. Failing to account for how parents and teachers adjust their behavior will lead researchers to incorrectly attribute differences in outcomes to differences in birth order (and ipso facto to differences in resources afforded to children of different parity).

While we find that, when one accounts for the indirect association between birth order and a variety of parental behavior, birth order differences persist on some outcomes (particularly on reading comprehension, reading recognition, and picture-vocabulary tests) the differences disappear for mathematics. Our results strongly suggest that parents and/or teachers adjust to remediate mathematics performance differences of children (who happen to be of higher order birth) but that any changes in the behavior of parents and/or teachers does not resolve differences on the other types of learning as measured by these tests.

Our results point to the many directions we will take in future research. We have not yet estimated the models allowing for state and family fixed effects. We will explore further the possible factors that might explain the gains in the percentile rank in the mathematics score distribution across time for children of higher order birth. Conversely, we will explore why there appears to be no closing of the gap on the other tests.

References

Adams, Scott J. (2002). "Educational attainment and health: Evidence from a sample of older adults." Education Economics 10, no. 1: 97-109.

Arendt, Jacob Nielsen. (2004). "Does education cause better health? A panel data analysis using school reforms for identification." Economics of Education Review 24, no. 2: 149-160.

Becker, Gary S. (1960). "An Economic Analysis of Fertility." In *Demographic and Economic Change in Developed Countries*. Universities-National Bureau Conference Series, no. 11. Princeton, NJ: Princeton University Press (for NBER).

Becker, Gary S. and Lewis, H. Gregg. "On the Interaction between the Quantity and Quality of Children." Journal of Political Economy, 81(2) pt. 2 (March/April):S279-S288.

Black, S., Devereux, P., Salvanes, K. (2005). "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education" Center for The Economics of Education, London School of Economics.

Center for Human Resource Research (2002). A Guide to the 1986–2000 Child Data 1994–2000 Young Adult Data. Ohio State University, Columbus, Ohio.

Chou, Shin-Yi, Jin-Tan Liu, Michael Grossman, and Theodore Joyce. (2004). "Parental education and child health: Evidence from a natural experiment in Taiwan." Working Paper, December.

Cicirelli. V.G. (1978). "The Relationship of Sibling Structure to Intellectual Abilities and Achievement," Review of Educational Research 48:365-369.

Crosser, S. (1991). "Summer Birth Date Children: Kindergarten Entrance Age and Academic Achievement." Journal of Educational Research, 84:140-146.

Cunha, F., Heckman, J.J., Lochner, L., and Masterov, D. (2005). "Interpreting the Evidence on Life Cycle Skill Formation" in E. Hanushek and F. Welch (eds) Handbook of the Economics of Education, North Holland, 2005.

Currie, J. and Thomas, D. (2000). "School Quality and the Longer-Term Effects of Head Start," Journal of Human Resources,35(4):755-774.

Datar, A. (2006) "The Impact of Kindergarten Entrance Age Policies on the Childcare Needs of Families.' Journal of Policy Analysis and Management, 25(1), 129-153.

Datar, A. (2003). The Impact of Changes in Kindergarten Entrance Age Policies on Academic Achievement and child Care Needs of Families. Dissertation, Rand Graduate School.

DeMeis, J.L., & Stearns, E.S. (1992). "Relationship of School Entrance Age to Academic and Social Performance." Journal of Educational Research, 86(1):20-27.

Fertig, Michael and Kluve, Jochen (2005). "The Effect of Age at School Entry on Educational Achievement in Germany." IZA discussion paper 1507.

Frederiksson, Peter and Öckert, Björn (2004). "Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labour Market Performance." Mimeo, University of Uppsala.

Gelbach, J.B. (2002). "Public Schooling for Young Children and Maternal Labor Supply." American Economic Review, 92(1), 307-322.

Gerner, Jennifer, and Lillard, Dean R.)2005). "Does School Performance Increase When Children Enter at Younger Ages?" Manuscript. Department of Policy Analysis and Management, Cornell University.

Gormley, W.T., Gayer, T., Phillips, D., and Dawson, B. (2004). "The Effects of Oklahoma's School Readiness Program on School Readiness: An Executive Summary." Center for Research on children in the U.S., Georgetown University.

Graue, M.E., & DiPerna, J. (2000). "Readshirting and Early Retention: Who Gets the "Gift of Time" and What are its Outcomes?" American Educational Research Journal, 37(2):509-534.

Hanushek, E.A. (1979). "Conceptual and Empirical Issues in the Estimation of Educational Production Functions." Journal of Human Resources 14 (Summer):351-88.

_______. (1992). "The Trade-off between Child Quantity and Quality," Journal of Political Economy, 100(1):84-117.

Hauser, R.M. and Sewell, W.H. (1985). "Birth Order and Educational Attainment in Full Sibships,"American Educational Research Journal, 22(1):1-23.

Heckman, James J. and Yona Rubinstein. (2001). "The importance of noncognitive skills: Lessons from the GED testing program." American Economic Review Papers and Proceedings 91, no.2: 145-149.

Jürges, Hendrik and Schneider, Kerstin. (2006). "Age at school entry and teacher's recommendations for secondary school track choice in Germany." Mimeo, University of Mannheim.

Kagan, S.L. (1990). "Readiness 2000: Rethinking rhetoric and responsibility." Phi Delta Kappan, December 1990: 272-279.

Lleras-Muney, Adriana. (2002). “Were compulsory attendance and child labor laws effective? An analysis from 1915 to 1939.” Journal of Law and Economics 45, no. 2: 401-435.

Schweinhart, L., Barnes, H. and Weikart, D (1993). Significant Benefits: The High-Scope Perry Pre-kindergarten Study through Age 27. Ypsilanti, Mich.: High Scope Press.

Shepard, Lorrie A. and Smith, Mary Lee. (1989). “Introduction and Overview” Chapter 1 in *Flunking Grades: Research and Policies on Retention*. Lorrie A. Shepard and Mary Lee Smith eds. New York: The Falmer Press.

Spasojevic, J. (2003). “Effects of Education on Adult Health in Sweden: Results from a Natural Experiment.” Ph.D. Dissertation (City University of New York Graduate Center, New York).

Steelman, L.C. (1985). “A Tale of Two Variables: A Review of the Intellectual Consequences of Sibship Size and Birth Order,” Review of Educational Research.

Stipek, D., & Byler, P. (2001). “Academic Achievement and Social Behaviors Associated with Age of Entry into Kindergarten.” Journal of Applied Developmental Psychology, 22:175-189.

U.S. Bureau of Economic Analysis “SA05 Personal Income by Major Source and Earnings by Industry” available at www.bea.gov///bea/regional/articles/spi2001/intro.cfm

U.S. Department of Labor. (1980-2002). *Geographic Profile of Employment and Unemployment*. Bureau of Labor Statistics, Bulletin 2361. Washington, DC (May).

Willis, Robert J. (1973). “A New Approach to the Economic Theory of Fertility Behavior.” Journal of Political Economy, 81 (2), pt. 2 (March/April): S14-S64.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: The MIT Press.

Zill, Nicholas, and West, Jerry. (2001). *Findings from the Condition of Education 2000: Entering Kindergarten*. U.S. Department of Education, Office of Education Research and Improvement. NCES 2001-035.

Table 1 Means and standard deviation of test score variables						
		Order of birth				
Variable	Full sample	1	2	3	4	5+
Mathematics						
Percentile score 1st test	50.66	53.55	50.87	46.93	42.95	36.04
	(27.65)	(26.92)	(28.06)	(27.28)	(27.19)	(27.93)
Age (in months) when took test 1	78.37	81.45	76.54	75.03	75.52	74.04
	(18.71)	(21.78)	(15.75)	(14.83)	(16.46)	(12.25)
Percentile score 2nd test	51.86	53.95	52.12	48.48	47.15	37.91
	(26.40)	(25.87)	(26.54)	(26.37)	(26.32)	(27.98)
Age (in months) when took test 2	104.03	107.58	101.85	99.95	98.80	100.50
	(19.65)	(22.54)	(16.53)	(16.22)	(14.31)	(14.42)
Percentile score 3rd test	53.33	55.75	53.47	49.41	47.47	37.15
	(28.51)	(27.92)	(28.48)	(28.44)	(29.86)	(30.26)
Age (in months) when took test 3	126.92	130.19	124.89	122.85	121.29	125.10
	(17.88)	(20.51)	(15.24)	(13.78)	(12.17)	(14.42)
PPVT						
Percentile score 1st test	36.59	42.12	35.94	28.96	23.76	20.28
	(29.54)	(30.00)	(29.06)	(27.07)	(25.32)	(24.63)
Age (in months) when took test 1	65.73	70.05	62.46	61.73	61.20	65.14
	(26.79)	(29.86)	(23.66)	(23.65)	(21.96)	(23.64)
Percentile score 2nd test	41.30	45.52	41.51	32.79	29.75	19.13
	(29.81)	(29.66)	(29.64)	(27.45)	(29.37)	(24.79)
Age (in months) when took test 2	116.62	120.45	113.65	112.78	109.67	115.13
	(26.60)	(29.22)	(24.08)	(22.59)	(21.92)	(22.52)
Percentile score 3rd test	45.22	49.67	43.96	35.81	31.58	30.67
	(29.78)	(30.24)	(28.86)	(26.84)	(27.27)	(30.43)
Age (in months) when took test 3	139.23	142.26	136.99	135.46	132.08	134.62
	(16.11)	(17.55)	(14.68)	(12.71)	(8.68)	(13.41)
Reading comprehension						
Percentile score 1st test	59.47	63.11	58.94	55.03	50.63	44.16
	(26.35)	(24.90)	(26.70)	(26.75)	(27.21)	(29.13)
Age (in months) when took test 1	92.54	93.34	91.36	91.88	92.66	97.05
	(18.50)	(20.63)	(16.26)	(16.40)	(18.03)	(16.92)
Percentile score 2nd test	54.58	58.50	54.01	49.44	43.22	34.33
	(28.39)	(27.67)	(28.02)	(28.63)	(27.76)	(29.06)
Age (in months) when took test 2	117.58	118.60	116.50	116.96	114.19	122.78
	(18.82)	(20.95)	(16.84)	(16.61)	(15.41)	(15.99)
Percentile score 3rd test	49.95	54.18	48.82	43.02	37.76	35.96
	(27.84)	(27.12)	(27.78)	(27.62)	(26.54)	(28.55)
Age (in months) when took test 3	139.48	140.48	138.47	138.55	137.29	142.79
	(16.97)	(19.00)	(15.21)	(14.14)	(13.67)	(13.64)
Reading recognition						
Percentile score 1st test	60.33	64.32	59.60	55.32	52.46	46.48
	(26.35)	(25.29)	(26.24)	(26.45)	(27.72)	(28.46)
Age (in months) when took test 1	78.76	81.90	76.89	75.34	76.10	74.06
	(18.63)	(21.69)	(15.71)	(14.55)	(16.61)	(12.38)
Percentile score 2nd test	58.49	61.94	58.34	53.79	48.84	41.91
	(27.70)	(26.53)	(27.84)	(28.55)	(27.20)	(29.12)
Age (in months) when took test 2	104.54	108.05	102.32	100.69	99.35	100.66
	(19.83)	(22.52)	(16.90)	(17.06)	(14.45)	(14.56)
Percentile score 3rd test	57.47	61.21	57.31	51.83	45.91	39.93
	(29.51)	(28.49)	(29.10)	(30.48)	(29.61)	(32.36)
Age (in months) when took test 3	127.53	130.86	125.38	123.52	122.05	125.41
	(18.21)	(20.81)	(15.43)	(14.65)	(12.42)	(14.61)

Source: authors' compilation from CNLSY79 and NLSY79. Sample sizes vary for each test. See Appendix Table 1.

Table 2 Sample statistics[1]

Variable	Mathematics		Reading recognition		Reading comprehension		PPVT	
Percentile ranking of test score (in test year)	48.13	(27.04)	58.08	(26.40)	57.70	(26.37)	33.40	(28.50)
Age when took test (years)	6.63	(1.65)	6.66	(1.65)	7.73	(1.59)	5.77	(2.44)
Birth order								
Born 1st	0.49	(0.50)	0.49	(0.50)	0.50	(0.50)	0.49	(0.50)
Born 2nd	0.32	(0.47)	0.32	(0.47)	0.32	(0.47)	0.32	(0.47)
Born 3rd	0.13	(0.34)	0.13	(0.34)	0.13	(0.33)	0.13	(0.34)
Born 4th	0.04	(0.20)	0.04	(0.20)	0.04	(0.19)	0.04	(0.19)
Born 5th or higher	0.02	(0.14)	0.02	(0.13)	0.02	(0.13)	0.02	(0.13)
Singleton	0.13	(0.34)	0.13	(0.34)	0.12	(0.32)	0.16	(0.36)
Youngest (last born)	0.42	(0.49)	0.42	(0.49)	0.39	(0.49)	0.46	(0.50)
Endogenous behaviors								
Age entered first grade (years)	6.30	(0.66)	6.30	(0.66)	6.31	(0.66)	6.29	(0.65)
Year began first grade	1989.98	(4.66)	1989.98	(4.67)	1989.75	(4.59)	1990.00	(4.63)
Attended preschool	0.59	(0.49)	0.59	(0.49)	0.59	(0.49)	0.59	(0.49)
Attended kindergarten	0.47	(0.50)	0.47	(0.50)	0.46	(0.50)	0.47	(0.50)
Years mother worked when child 0-5	0.31	(0.31)	0.32	(0.31)	0.32	(0.31)	0.32	(0.31)
Mother began working in year entered first grade	0.13	(0.34)	0.13	(0.34)	0.13	(0.34)	0.14	(0.34)
Other covariates								
Adults in household	1.93	(0.74)	1.92	(0.74)	1.90	(0.71)	1.96	(0.74)
Youth <18 in household	2.48	(1.14)	2.48	(1.14)	2.53	(1.15)	2.38	(1.11)
Number of siblings	1.62	(1.17)	1.62	(1.17)	1.67	(1.20)	1.52	(1.16)
Multiple birth	0.02	(0.13)	0.02	(0.13)	0.02	(0.12)	0.02	(0.12)
Unanticipated increase in family size	0.02	(0.13)	0.02	(0.13)	0.02	(0.12)	0.01	(0.12)
Black	0.31	(0.46)	0.31	(0.46)	0.31	(0.46)	0.30	(0.46)
Hispanic	0.20	(0.40)	0.20	(0.40)	0.20	(0.40)	0.19	(0.40)
Female	0.51	(0.50)	0.51	(0.50)	0.50	(0.50)	0.51	(0.50)
Mother's age	31.17	(3.93)	31.20	(3.94)	32.09	(4.14)	30.35	(3.72)
Mother's education	12.22	(2.21)	12.22	(2.21)	12.24	(2.22)	12.23	(2.17)
Mother's Rotter index score	9.33	(2.24)	9.34	(2.24)	9.34	(2.23)	9.30	(2.23)
Maternal grandmother's education	9.60	(3.97)	9.60	(3.97)	9.57	(3.98)	9.66	(3.94)
Average family income ('0,000s of 2002 $)	8.70	(11.79)	8.70	(11.74)	8.89	(12.27)	8.66	(11.53)
Mother married	0.43	(0.50)	0.43	(0.50)	0.42	(0.49)	0.44	(0.50)
Mother separated/widowed/divorced	0.25	(0.43)	0.25	(0.43)	0.26	(0.44)	0.23	(0.42)
Compulsory schooling laws								
Statutory age permitted to enter school (years)	6.37	(0.75)	6.37	(0.75)	6.37	(0.75)	6.37	(0.75)
Statutory age must enter school (years)	5.42	(0.83)	5.42	(0.83)	5.42	(0.83)	5.41	(0.82)
Permitted age imputed	0.16	(0.37)	0.00	(0.04)	0.00	(0.00)	0.15	(0.36)
Log of state's per capita earnings in:								
Agriculture	-2.26	(0.46)	-2.26	(0.46)	-2.27	(0.46)	-2.26	(0.46)
Mining	-2.19	(1.15)	-2.19	(1.15)	-2.18	(1.15)	-2.20	(1.15)
Construction	0.11	(0.22)	0.11	(0.22)	0.10	(0.22)	0.11	(0.22)
Manufacturing	1.30	(0.38)	1.30	(0.38)	1.30	(0.38)	1.30	(0.38)
Transportation and public utilities	0.27	(0.21)	0.27	(0.21)	0.27	(0.21)	0.27	(0.21)
Wholesale trade	0.16	(0.27)	0.16	(0.27)	0.16	(0.27)	0.16	(0.27)
Retail trade	0.60	(0.13)	0.60	(0.13)	0.60	(0.13)	0.60	(0.13)
Finance, insurance, and real estate	0.21	(0.46)	0.21	(0.46)	0.20	(0.46)	0.21	(0.46)
Services	1.50	(0.34)	1.50	(0.34)	1.49	(0.34)	1.50	(0.34)
Civilian employment for federal government	-0.40	(0.42)	-0.40	(0.42)	-0.40	(0.43)	-0.41	(0.43)
Military	-1.24	(0.85)	-1.24	(0.85)	-1.23	(0.85)	-1.25	(0.85)
State and local government	0.84	(0.18)	0.84	(0.18)	0.84	(0.18)	0.84	(0.18)
State unemployment rate	6.45	(1.91)	6.45	(1.91)	6.51	(1.92)	6.45	(1.90)
N	5629	5629	5625	5625	5281	5281	5508	5508

[1]Means taken for year child took first mathematics test except where obviously linked to other year or where time-invariant.
Source: Children of the NLSY79, NLSY79, Compulsory Schooling Law (CoSLaw) data base (compiled by authors), and Bureau of Economic Analysis series SA05 "Personal Income by Major Source and Earnings by Industry."

Table 3 Sample statistics for respondents with siblings who faced different statutory schooling ages, by whether respondent entered school in same/different state than the average state of all siblings

	Same state			Different state		
Statutory age	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Permitted age	5.44	(0.83)	3772	5.34	(0.81)	546
Compulsory age	6.37	(0.75)	3772	6.43	(0.77)	546

	Same state		Different state		Combined sample	
Permitted age faced by respondent was:	Percent	N	Percent	N	Percent	N
less than avg. permitted age faced by all siblings	8.01%	346	5.23%	226	13.25%	572
same as avg. permitted age faced by all siblings	70.03%	3024	2.15%	93	72.19%	3117
older than avg. permitted age faced by all siblings	9.31%	402	5.26%	227	14.57%	629

	Same state		Different state		Combined sample	
Compulsory age faced by respondent was:	Percent	N	Percent	N	Percent	N
less than avg. compulsory age faced by all siblings	8.89%	384	5.49%	237	14.38%	621
same as avg. compulsory age faced by all siblings	68.46%	2956	1.48%	64	69.94%	3020
older than avg. compulsory age faced by all siblings	10.00%	432	5.67%	245	15.68%	677

Notes: Averages are calculated over all respondents with 1 or more siblings with a valid score on the first math test.
Source: Authors' calculations from Children of NLSY79 and Compulsory Schooling Law (CoSLaw) data base (compiled by authors).

Table 4 Regression estimates of birth order penalty on mathematics test (in percentile score rank points)

	Test 1			Test 2			Test 3		
Variable	1	2	3	1	2	3	1	2	3
Birth order penalty (relative to first born)									
Singleton	-1.76	1.92	0.00	-1.08	2.60	-0.13	0.12	3.90 *	3.37 *
	(1.83)	(1.71)	(1.43)	(1.84)	(1.69)	(1.48)	(2.18)	(2.00)	(1.92)
Born 2nd	-2.77 ***	-1.69 *	-2.60 **	-2.03 **	-0.55	-1.12	-2.41 **	-0.44	-0.44
	(1.01)	(0.94)	(1.15)	(1.02)	(0.93)	(1.07)	(1.20)	(1.10)	(1.30)
Born 3rd	-5.23 ***	-2.64 **	-3.29 *	-3.75 ***	-0.63	-0.94	-4.62 ***	-0.56	-0.13
	(1.35)	(1.26)	(1.80)	(1.37)	(1.26)	(1.66)	(1.62)	(1.49)	(2.00)
Born 4th	-7.12 ***	-3.41 *	-3.03	-2.22	2.23	-1.46	-3.97	3.18	1.75
	(2.16)	(2.01)	(2.70)	(2.24)	(2.05)	(2.51)	(2.73)	(2.52)	(3.06)
Born 5th or later	-11.57 ***	-5.39 **	-10.58 ***	-7.54 **	-0.78	-3.86	-9.18 **	-1.14	-3.26
	(2.87)	(2.68)	(3.78)	(3.05)	(2.80)	(3.50)	(3.72)	(3.42)	(4.41)
Control variables									
Number of siblings	yes	yes	yes	yes	yes	yes	yes	yes	yes
Age took test	yes	yes	yes	yes	yes	yes	yes	yes	yes
Demographics	no	yes	yes	no	yes	yes	no	yes	yes
Endogenous variables controlled?									
Attended pre-kindergarten	no	no	yes	no	no	yes	no	no	yes
Attended kindergarten	no	no	yes	no	no	yes	no	no	yes
Age entered 1st grade	no	no	yes	no	no	yes	no	no	yes
Years mother worked when child was age 0-5	no	no	yes	no	no	yes	no	no	yes
Entered labor force year entered 1st grade	no	no	yes	no	no	yes	no	no	yes
N	4209	4209	5525	3691	3691	4776	3079	3079	3667
R-Square	0.033	0.165	0.171	0.054	0.210	0.225	0.060	0.217	0.112
Dep. Mean	50.66	50.66		51.88	51.88		53.32	53.32	

Notes: Coefficients that are statistically different from zero are denoted by *** for p≤.001, ** for p≤.05 and * for p≤.10. Column 1 controls for number of siblings and age when tested, column 2 adds race indicators for black and other race, mother's highest grade completed, maternal grandmother's highest grade completed, and the mother's Rotter index score. Column 3 estimates effect of birth order on test scores using three-stage least squares (simultaneous equation) models that include: whether a child attended pre-kindergarten schooling, kindergarten, the age he enrolled in first grade, the number of years his mother worked in the first six years of his life, and whether his mother entered the labor force in the year he entered first grade.

Table 5 Regression estimates of birth order penalty on Peabody Picture-Vocabulary Test (in percentile score rank points)

	Test 1			Test 2			Test 3		
Variable	1	2	3	1	2	3	1	2	3
Birth order penalty (relative to first born)									
Singleton	-6.40 ***	-1.32	-0.84	-1.43	2.61	-1.20	2.11	5.36 *	0.79
	(1.91)	(1.66)	(1.43)	(2.33)	(2.06)	(1.74)	(3.15)	(2.80)	(2.06)
Born 2nd	-7.67 ***	-5.87 ***	-1.18	-4.53 ***	-2.85 **	-4.83 ***	-5.98 ***	-4.57 ***	-3.98 ***
	(1.07)	(0.93)	(1.11)	(1.27)	(1.12)	(1.21)	(1.70)	(1.51)	(1.34)
Born 3rd	-13.19 ***	-9.13 ***	-0.33	-11.77 ***	-7.46 ***	-6.08 ***	-13.40 ***	-8.95 ***	-7.95 ***
	(1.41)	(1.23)	(1.73)	(1.74)	(1.54)	(1.88)	(2.49)	(2.23)	(2.09)
Born 4th	-16.45 ***	-10.66 ***	-0.23	-13.71 ***	-7.19 ***	-7.21 **	-16.82 ***	-9.72 ***	-11.23 ***
	(2.27)	(1.98)	(2.59)	(3.05)	(2.70)	(2.88)	(4.18)	(3.73)	(3.25)
Born 5th or later	-16.59 ***	-7.47 ***	2.18	-19.84 ***	-10.59 ***	-8.85 **	-14.29 **	-10.49 *	-10.49 **
	(3.00)	(2.62)	(3.64)	(4.12)	(3.66)	(4.16)	(6.99)	(6.21)	(4.95)
Control variables									
Number of siblings	yes	yes	yes	yes	yes	yes	yes	yes	yes
Age took test	yes	yes	yes	yes	yes	yes	yes	yes	yes
Demographics	no	yes	yes	no	yes	yes	no	yes	yes
Endogenous variables controlled?									
Attended pre-kindergarten	no	no	yes	no	no	yes	no	no	yes
Attended kindergarten	no	no	yes	no	no	yes	no	no	yes
Age entered 1st grade	no	no	yes	no	no	yes	no	no	yes
Years mother worked when child was age 0-5	no	no	yes	no	no	yes	no	no	yes
Entered labor force year entered 1st grade	no	no	yes	no	no	yes	no	no	yes
N	4182	4182	5162	2906	2906	4148	1603	1603	3107
R-Square	0.073	0.302	-0.250	0.070	0.278	0.212	0.056	0.262	0.235
Dep. Mean	36.65	36.65		41.27	41.27		45.22	45.22	

Notes: Coefficients that are statistically different from zero are denoted by *** for p≤.001, ** for p≤.05 and * for p≤.10. Column 1 controls for number of siblings and age when tested, column 2 adds race indicators for black and other race, mother's highest grade completed, maternal grandmother's highest grade completed, and the mother's Rotter index score. Column 3 estimates effect of birth order on test scores using three-stage least squares (simultaneous equation) models that include: whether a child attended pre-kindergarten schooling, kindergarten, the age he enrolled in first grade, the number of years his mother worked in the first six years of his life, and whether his mother entered the labor force in the year he entered first grade.

Table 6 Regression estimates of birth order penalty on reading comprehension test (in percentile score rank points)

	Test 1			Test 2			Test 3		
Variable	1	2	3	1	2	3	1	2	3
Birth order penalty (relative to first born)									
Singleton	-0.46	1.44	-0.70	-1.77	0.96	1.67	-3.63	-0.31	-0.59
	(1.63)	(1.59)	(1.36)	(1.93)	(1.84)	(1.86)	(2.21)	(2.06)	(2.67)
Born 2nd	-4.72 ***	-4.27 ***	-5.07 ***	-4.97 ***	-4.41 ***	-4.49 ***	-5.81 ***	-4.75 ***	-5.01 ***
	(0.90)	(0.87)	(1.19)	(1.06)	(1.00)	(1.29)	(1.19)	(1.10)	(1.83)
Born 3rd	-6.99 ***	-5.81 ***	-7.89 ***	-7.51 ***	-5.68 ***	-8.04 ***	-9.69 ***	-7.20 ***	-9.58 ***
	(1.20)	(1.16)	(1.86)	(1.44)	(1.36)	(2.01)	(1.65)	(1.54)	(2.83)
Born 4th	-9.30 ***	-7.30 ***	-8.14 ***	-13.17 ***	-9.42 ***	-6.72 **	-13.07 ***	-7.80 ***	-8.24 *
	(1.92)	(1.86)	(2.77)	(2.41)	(2.29)	(3.13)	(2.76)	(2.58)	(4.38)
Born 5th or later	-11.13 ***	-8.40 ***	-3.28	-13.94 ***	-9.59 ***	-9.02 **	-8.61 **	-4.10	-5.39
	(2.64)	(2.56)	(3.82)	(3.39)	(3.22)	(4.49)	(4.13)	(3.85)	(7.91)
Control variables									
Number of siblings	yes	yes	yes	yes	yes	yes	yes	yes	yes
Age took test	yes	yes	yes	yes	yes	yes	yes	yes	yes
Demographics	no	yes	yes	no	yes	yes	no	yes	yes
Endogenous variables controlled?									
Attended pre-kindergarten	no	no	yes	no	no	yes	no	no	yes
Attended kindergarten	no	no	yes	no	no	yes	no	no	yes
Age entered 1st grade	no	no	yes	no	no	yes	no	no	yes
Years mother worked when child was age 0-5	no	no	yes	no	no	yes	no	no	yes
Entered labor force year entered 1st grade	no	no	yes	no	no	yes	no	no	yes
N	3971	3971	5433	3350	3350	3700	2659	2659	1981
R-Square	0.194	0.247	0.303	0.173	0.260	0.273	0.133	0.254	0.247
Dep. Mean	59.48	59.48		54.60	54.60		49.94	49.94	

Notes: Coefficients that are statistically different from zero are denoted by *** for p≤.001, ** for p≤.05 and * for p≤.10. Column 1 controls for number of siblings and age when tested, column 2 adds race indicators for black and other race, mother's highest grade completed, maternal grandmother's highest grade completed, and the mother's Rotter index score. Column 3 estimates effect of birth order on test scores using three-stage least squares (simultaneous equation) models that include: whether a child attended pre-kindergarten schooling, kindergarten, the age he enrolled in first grade, the number of years his mother worked in the first six years of his life, and whether his mother entered the labor force in the year he entered first grade.

Table 7 Regression estimates of birth order penalty on reading recognition test (in percentile score rank points)

	Test 1			Test 2			Test 3		
Variable	1	2	3	1	2	3	1	2	3
Birth order penalty (relative to first born)									
Singleton	-1.48	0.66	-0.42	-1.11	1.92	0.01	-1.87	1.19	0.35
	(1.71)	(1.63)	(1.37)	(1.91)	(1.80)	(1.60)	(2.26)	(2.11)	(2.04)
Born 2nd	-5.17 ***	-4.36 ***	-5.75 ***	-3.99 ***	-2.71 ***	-2.44 **	-4.41 ***	-2.73 **	-4.33 ***
	(0.95)	(0.90)	(1.11)	(1.06)	(0.99)	(1.16)	(1.23)	(1.15)	(1.37)
Born 3rd	-7.69 ***	-5.71 ***	-8.84 ***	-6.72 ***	-3.89 ***	-4.34 **	-7.97 ***	-4.37 ***	-6.93 ***
	(1.26)	(1.20)	(1.73)	(1.42)	(1.34)	(1.80)	(1.67)	(1.57)	(2.11)
Born 4th	-7.73 ***	-5.18 ***	-7.43 ***	-9.45 ***	-5.40 **	-7.49 ***	-11.29 ***	-5.03 *	-6.54 **
	(2.01)	(1.92)	(2.60)	(2.33)	(2.19)	(2.72)	(2.82)	(2.64)	(3.23)
Born 5th or later	-10.79 ***	-6.65 ***	-10.24 ***	-12.13 ***	-6.28 **	-5.50	-12.30 ***	-5.11	-9.36 **
	(2.68)	(2.56)	(3.63)	(3.15)	(2.97)	(3.81)	(3.84)	(3.59)	(4.65)
Control variables									
Number of siblings	yes	yes	yes	yes	yes	yes	yes	yes	yes
Age took test	yes	yes	yes	yes	yes	yes	yes	yes	yes
Demographics	no	yes	yes	no	yes	yes	no	yes	yes
Endogenous variables controlled?									
Attended pre-kindergarten	no	no	yes	no	no	yes	no	no	yes
Attended kindergarten	no	no	yes	no	no	yes	no	no	yes
Age entered 1st grade	no	no	yes	no	no	yes	no	no	yes
Years mother worked when child was age 0-5	no	no	yes	no	no	yes	no	no	yes
Entered labor force year entered 1st grade	no	no	yes	no	no	yes	no	no	yes
N	4204	4204	5517	3688	3688	4759	3072	3072	3641
R-Square	0.072	0.162	0.1604	0.072	0.184	0.1578	0.067	0.191	0.163
Dep. Mean	60.35	60.35		58.50	58.50		57.46	57.46	

Notes: Coefficients that are statistically different from zero are denoted by *** for p≤.001, ** for p≤.05 and * for p≤.10. Column 1 controls for number of siblings and age when tested, column 2 adds race indicators for black and other race, mother's highest grade completed, maternal grandmother's highest grade completed, and the mother's Rotter index score. Column 3 estimates effect of birth order on test scores using three-stage least squares (simultaneous equation) models that include: whether a child attended pre-kindergarten schooling, kindergarten, the age he enrolled in first grade, the number of years his mother worked in the first six years of his life, and whether his mother entered the labor force in the year he entered first grade.

Table 8 Gain in percentile test score between 1st and 2nd standardized test

	Mathematics						Peabody Picture-Vocabulary Test					
Variable	1		2		3		1		2		3	
Birth order penalty												
Singleton	2.48		0.51		0.77		3.05	*	2.00		1.81	
	(1.65)		(1.86)		(1.87)		(1.79)		(2.03)		(2.03)	
Born 2nd	1.40		2.09	*	2.17	*	3.73	***	4.09	***	4.09	***
	(1.12)		(1.16)		(1.16)		(1.19)		(1.25)		(1.25)	
Born 3rd	0.19		1.88		2.03		4.22	***	4.99	***	4.94	***
	(1.45)		(1.63)		(1.63)		(1.58)		(1.78)		(1.79)	
Born 4th	4.23	*	6.91	***	7.17	***	5.43	**	6.79	**	6.75	**
	(2.16)		(2.47)		(2.47)		(2.57)		(2.88)		(2.89)	
Born 5th or higher	3.26		7.03	**	7.51	**	-1.46		0.23		0.09	
	(2.76)		(3.21)		(3.21)		(3.23)		(3.72)		(3.73)	
Last born	2.28	**	0.76		0.75		-0.69		-1.42		-1.39	
	(1.14)		(1.32)		(1.32)		(1.26)		(1.45)		(1.45)	
N	3688		3688		3688		2901		2901		2901	
R-Square	0.004		0.006		0.008		0.006		0.007		0.010	
Dep_Mean	1.56		1.56		1.56		5.28		5.28		5.28	
	Reading comprehension						Reading recognition					
Variable	1		2		3		1		2		3	
Birth order penalty												
Singleton	-0.68		-2.15		-0.95		1.15		0.64		1.82	
	(1.61)		(1.80)		(1.79)		(1.55)		(1.75)		(1.74)	
Born 2nd	-0.38		-0.11		-0.02		2.35	**	2.45	**	2.59	**
	(1.09)		(1.13)		(1.12)		(1.05)		(1.09)		(1.09)	
Born 3rd	-1.01		-0.09		0.47		1.55		1.97		2.47	
	(1.43)		(1.59)		(1.58)		(1.36)		(1.53)		(1.52)	
Born 4th	-5.29	**	-4.26	*	-3.31		-0.68		-0.05		0.58	
	(2.19)		(2.47)		(2.46)		(2.04)		(2.32)		(2.31)	
Born 5th or higher	-4.87	*	-2.30		-1.26		-0.23		0.98		1.89	
	(2.90)		(3.29)		(3.27)		(2.58)		(3.00)		(2.98)	
Last born	1.42		0.18		0.21		0.42		-0.17		-0.19	
	(1.13)		(1.30)		(1.28)		(1.07)		(1.23)		(1.22)	
N	3345		3345		3345		3685		3685		3685	
R-Square	0.003		0.025		0.044		0.002		0.008		0.026	
Mean change in test score	-5.19		-5.19		-5.19		-1.58		-1.58		-1.58	
Control variables												
Age took test	No		Yes		Yes		No		Yes		Yes	
Number of siblings	No		Yes		Yes		No		Yes		Yes	
Demographics	No		No		Yes		No		No		Yes	

Notes: Dependent variable is percentile score on 2nd standarized test minus percentile score on 1st standardized test.
Age and number of siblings are measured as the difference in the variable between the years the child took each test. Demographic variables (in levels) include race (black, other race), mother's education, maternal grandmother's education, and the mother's Rotter index score. Coefficients that are statistically different from zero are denoted by *** for p≤.001, ** for p≤.05 and * for p≤.10.

Table 9 Gain in percentile test score between 1st and 3rd standardized test

	Mathematics						Peabody Picture-Vocabulary Test					
Variable	1		2		3		1		2		3	
Birth order penalty												
Singleton	4.19	**	1.12		1.66		2.93		3.79		3.53	
	(1.98)		(2.23)		(2.23)		(2.47)		(2.83)		(2.85)	
Born 2nd	2.55	*	3.65	***	3.80	***	3.49	**	3.15	*	3.17	*
	(1.32)		(1.37)		(1.37)		(1.65)		(1.75)		(1.75)	
Born 3rd	1.02		3.65	*	4.10	**	4.83	**	4.01		3.92	
	(1.72)		(1.95)		(1.95)		(2.33)		(2.65)		(2.65)	
Born 4th	3.76		7.99	***	8.87	***	3.65		2.41		2.25	
	(2.67)		(3.05)		(3.05)		(3.53)		(4.04)		(4.05)	
Born 5th or higher	2.66		8.76	**	9.75	**	-1.00		-3.08		-2.84	
	(3.40)		(3.97)		(3.97)		(5.66)		(6.52)		(6.54)	
Last born	1.42		-0.96		-0.89		0.80		1.50		1.61	
	(1.37)		(1.59)		(1.59)		(1.80)		(2.10)		(2.11)	
N	3078		3078		3078		1602		1602		1602	
R-Square	0.004		0.007		0.013		0.007		0.007		0.010	
Dep_Mean	3.30		3.30		3.30		7.50		7.50		7.50	
	Reading comprehension						Reading recognition					
Variable	1		2		3		1		2		3	
Birth order penalty												
Singleton	-0.84		-1.16		0.51		1.36		0.00		1.24	
	(1.94)		(2.18)		(2.15)		(1.91)		(2.15)		(2.13)	
Born 2nd	-1.18		-1.35		-1.02		1.98		2.33	*	2.51	*
	(1.27)		(1.33)		(1.31)		(1.27)		(1.32)		(1.30)	
Born 3rd	-1.84		-1.90		-0.95		1.49		2.54		3.32	*
	(1.70)		(1.91)		(1.88)		(1.65)		(1.87)		(1.85)	
Born 4th	-8.23	***	-8.10	***	-6.36	**	-2.54		-0.87		0.40	
	(2.63)		(2.97)		(2.93)		(2.57)		(2.92)		(2.90)	
Born 5th or higher	-1.74		-1.43		0.01		-0.94		2.04		3.28	
	(3.73)		(4.24)		(4.17)		(3.27)		(3.81)		(3.77)	
Last born	1.65		1.36		1.16		0.31		-0.89		-0.93	
	(1.37)		(1.57)		(1.55)		(1.32)		(1.52)		(1.51)	
N	2657		2657		2657		3071		3071		3071	
R-Square	0.004		0.015		0.051		0.002		0.010		0.036	
Mean change in test score	-10.56		-10.56		-10.56		-2.55		-2.55		-2.55	
Control variables												
Age took test	No		Yes		Yes		No		Yes		Yes	
Number of siblings	No		Yes		Yes		No		Yes		Yes	
Demographics	No		No		Yes		No		No		Yes	

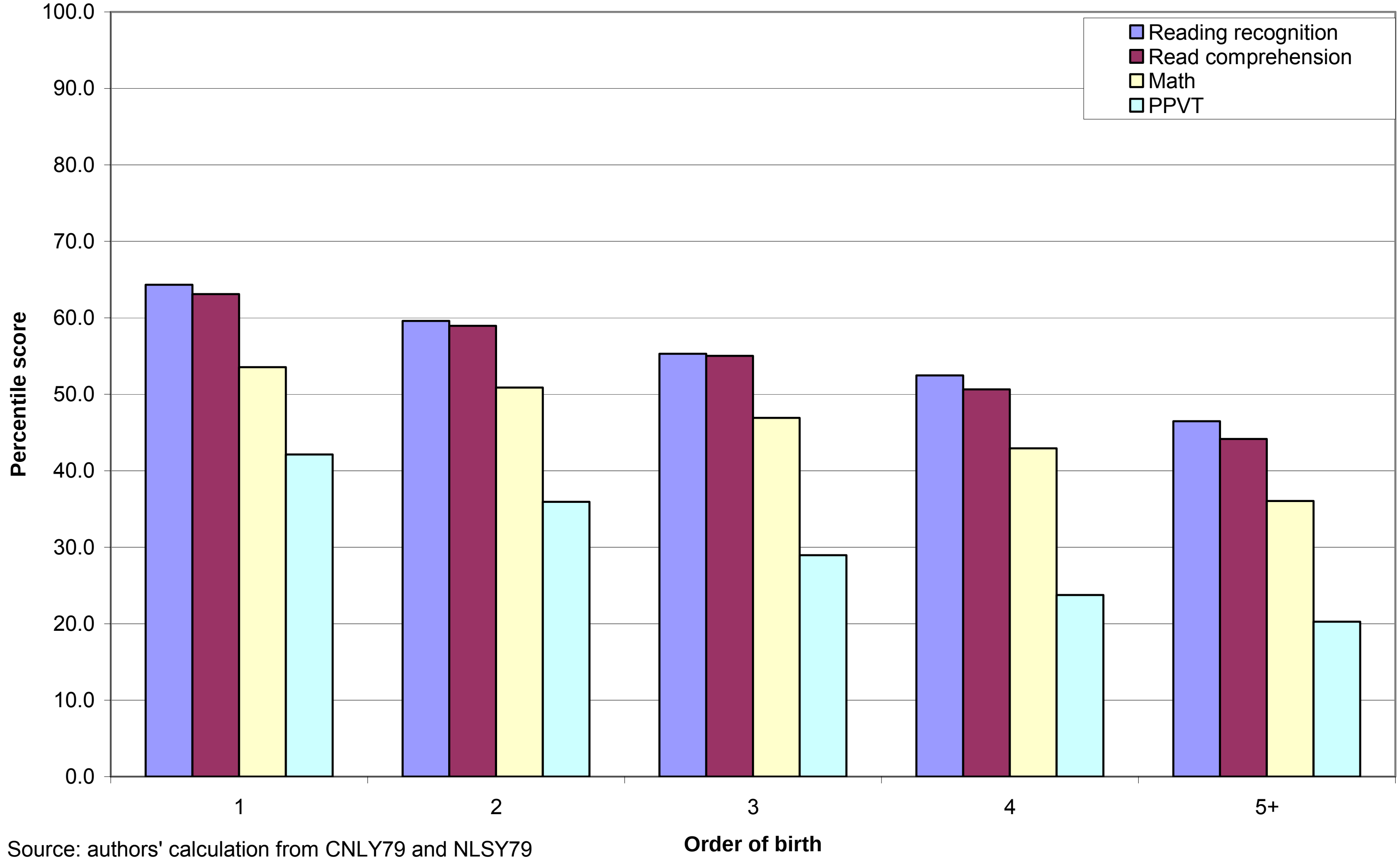
Notes: Dependent variable is percentile score on 2nd standarized test minus percentile score on 1st standardized test.
Age and number of siblings are measured as the difference in the variable between the years the child took each test. Demographic variables (in levels) include race (black, other race), mother's education, maternal grandmother's education, and the mother's Rotter index score. Coefficients that are statistically different from zero are denoted by *** for p≤.001, ** for p≤.05 and * for p≤.10.

Table 10 Gain in percentile test score between 2nd and 3rd standardized test

	Mathematics			Peabody Picture-Vocabulary Test		
Variable	1	2	3	1	2	3
Birth order penalty						
Singleton	1.08	0.28	0.67	-0.99	-1.43	-1.15
	(1.59)	(1.80)	(1.81)	(2.19)	(2.52)	(2.53)
Born 2nd	0.66	0.97	1.05	0.83	1.00	1.13
	(1.06)	(1.11)	(1.11)	(1.47)	(1.55)	(1.55)
Born 3rd	0.70	1.39	1.72	0.92	1.28	1.67
	(1.39)	(1.57)	(1.58)	(2.07)	(2.35)	(2.36)
Born 4th	-0.19	0.92	1.51	-3.58	-3.02	-2.39
	(2.16)	(2.47)	(2.47)	(3.14)	(3.59)	(3.59)
Born 5th or higher	0.03	1.55	2.21	4.47	5.41	6.12
	(2.75)	(3.21)	(3.21)	(5.03)	(5.80)	(5.80)
Last born	-0.49	-1.11	-1.09	0.62	0.28	0.36
	(1.11)	(1.28)	(1.28)	(1.60)	(1.87)	(1.87)
N	3075	3075	3075	1602	1602	1602
R-Square	0.000	0.001	0.005	0.003	0.003	0.009
Dep_Mean	1.77	1.77	1.77	1.90	1.90	1.90
	Reading comprehension			Reading recognition		
Variable	1	2	3	1	2	3
Birth order penalty						
Singleton	-1.66	-1.62	-1.22	-1.04	-1.98	-1.52
	(1.73)	(1.95)	(1.95)	(1.23)	(1.38)	(1.38)
Born 2nd	-1.50	-1.54	-1.47	-0.61	-0.24	-0.16
	(1.13)	(1.19)	(1.19)	(0.82)	(0.85)	(0.85)
Born 3rd	-1.90	-1.85	-1.56	-0.42	0.45	0.80
	(1.52)	(1.71)	(1.71)	(1.06)	(1.20)	(1.20)
Born 4th	-0.94	-0.86	-0.39	-0.49	0.86	1.45
	(2.35)	(2.66)	(2.66)	(1.66)	(1.89)	(1.89)
Born 5th or higher	2.63	2.76	3.16	0.31	2.31	2.98
	(3.32)	(3.78)	(3.78)	(2.11)	(2.46)	(2.45)
Last born	0.39	0.40	0.35	0.02	-0.72	-0.74
	(1.22)	(1.41)	(1.40)	(0.85)	(0.98)	(0.98)
N	2656	2656	2656	3068	3068	3068
R-Square	0.002	0.008	0.013	0.000	0.002	0.012
Mean change in test score	-5.65	-5.65	-5.65	-0.89	-0.89	-0.89
Control variables						
Age took test	No	Yes	Yes	No	Yes	Yes
Number of siblings	No	Yes	Yes	No	Yes	Yes
Demographics	No	No	Yes	No	No	Yes

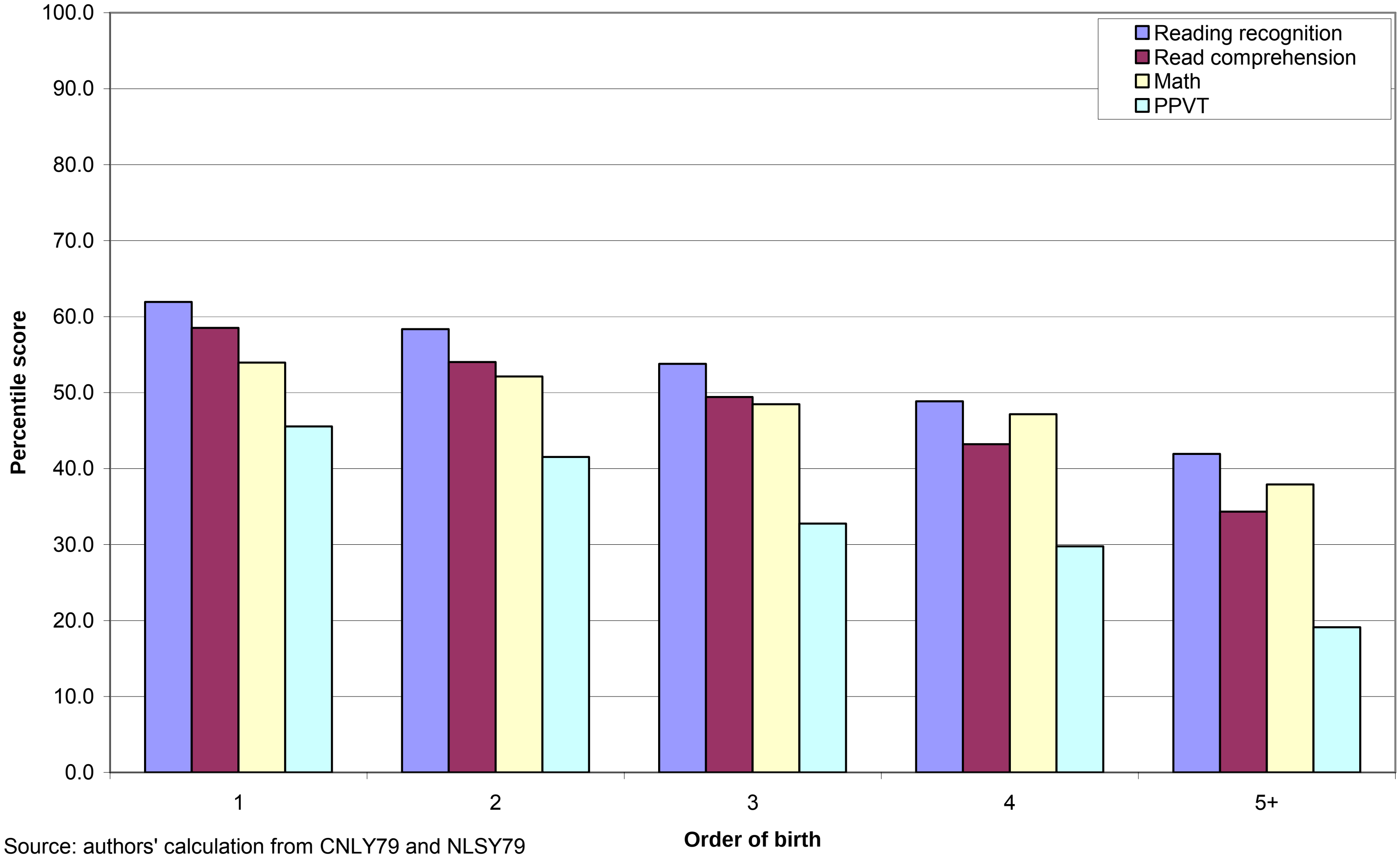
Notes: Dependent variable is percentile score on 2nd standarized test minus percentile score on 1st standardized test.
Age and number of siblings are measured as the difference in the variable between the years the child took each test. Demographic variables (in levels) include race (black, other race), mother's education, maternal grandmother's education, and the mother's Rotter index score. Coefficients that are statistically different from zero are denoted by *** for p≤.001, ** for p≤.05 and * for p≤.10.

Figure 1 Percentile scores on 1st standardized test, by birth order



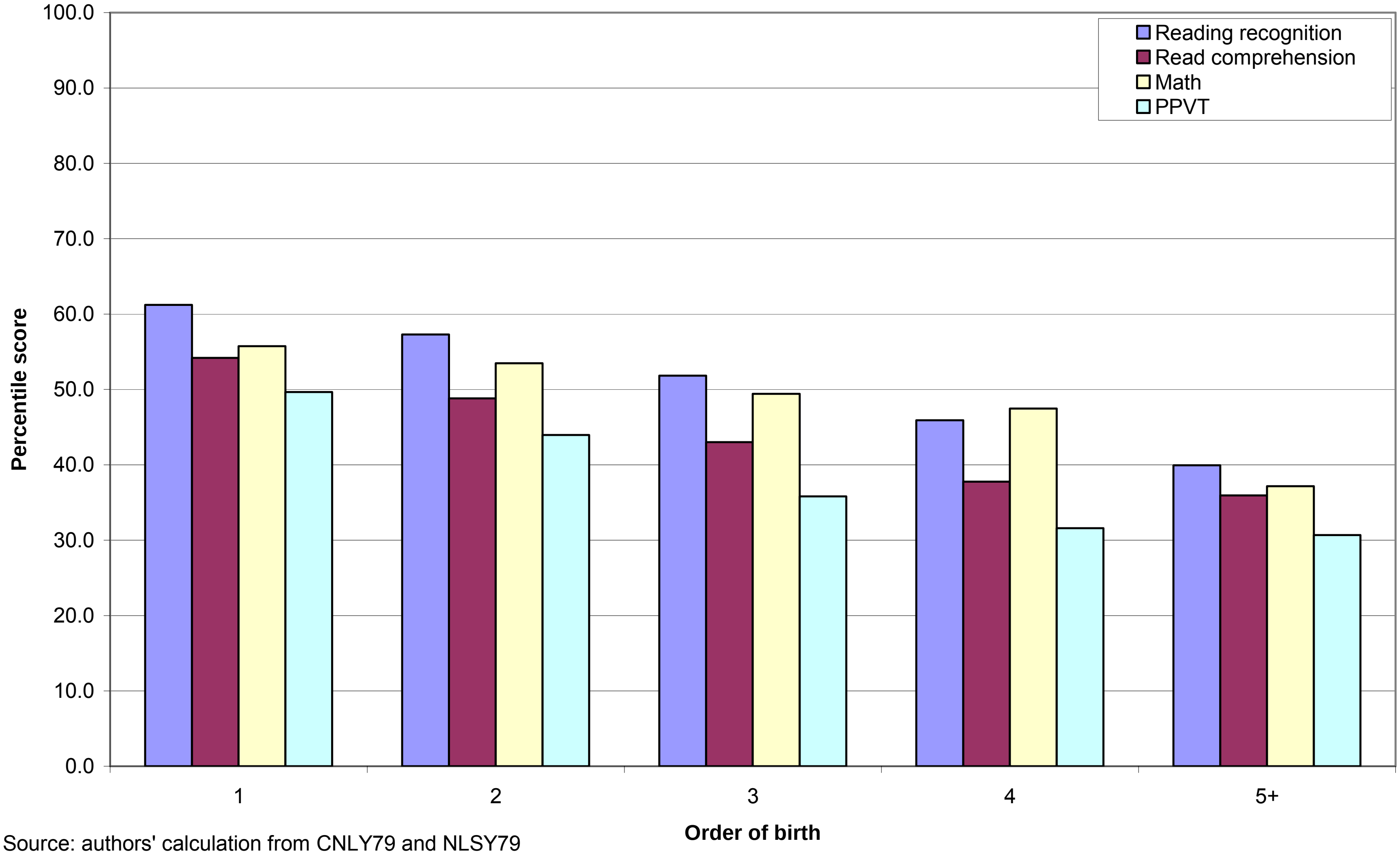
Source: authors' calculation from CNLY79 and NLSY79

Figure 2 Percentile scores on 2nd standardized test, by birth order



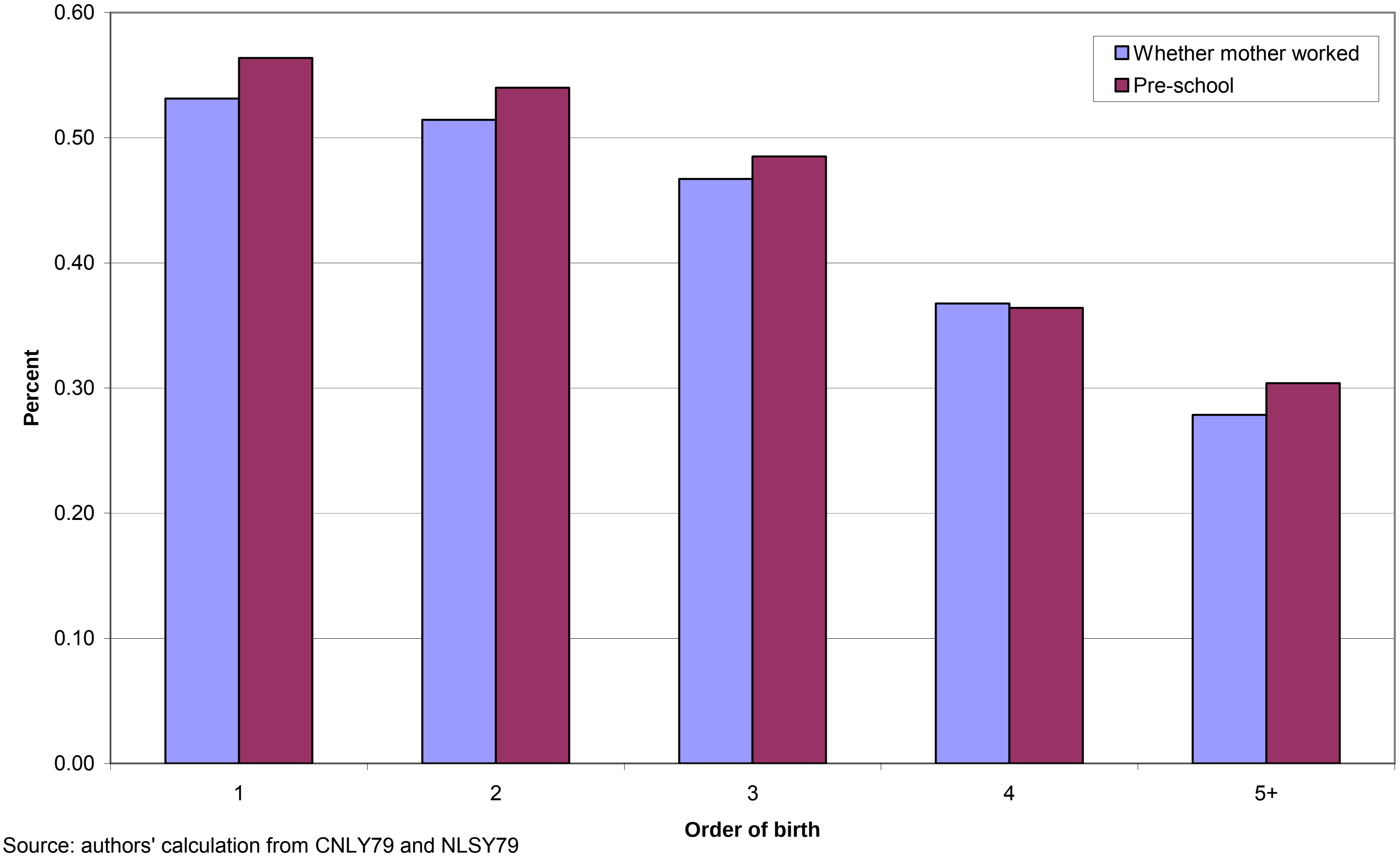
Source: authors' calculation from CNLY79 and NLSY79

Figure 3 Percentile scores on 3rd standardized test, by birth order



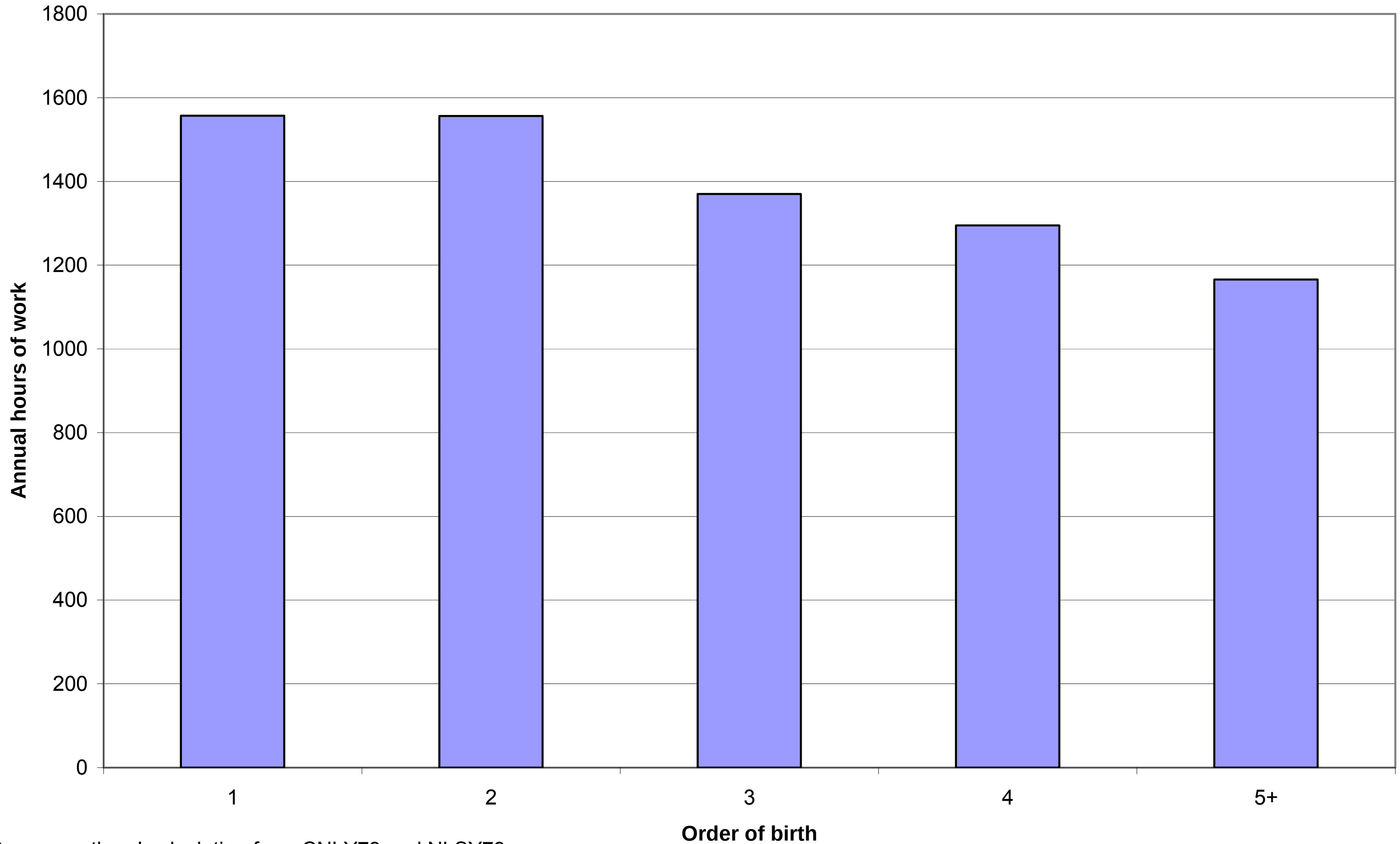
Source: authors' calculation from CNLY79 and NLSY79

Figure 4 Enrollment in pre-school and mother's work status, by birth order



Source: authors' calculation from CNLY79 and NLSY79

Figure 5 Mother's hours of paid market work in year child entered school, by birth order



Source: authors' calculation from CNLY79 and NLSY79

Appendix Table 1 Sample sizes associated with each standardized test score						
		Order of birth				
Variable	Full sample	1	2	3	4	5
PIAT Math test 1	4209	1863	1395	626	205	120
PIAT Math test 2	3691	1695	1203	532	169	92
PIAT Math test 3	3079	1434	1009	436	129	71
PPVT test 1	4182	1852	1381	626	202	121
PPVT test 2	2906	1372	965	400	108	61
PPVT test 3	1603	802	532	188	60	21
PIAT Reading comprehension test 1	3971	1791	1308	577	192	103
PIAT Reading comprehension test 2	3350	1563	1102	470	143	72
PIAT Reading comprehension test 3	2659	1283	867	352	109	48
PIAT Reading recognition test 1	4204	1858	1396	625	205	120
PIAT Reading recognition test 2	3688	1693	1203	531	168	93
PIAT Reading recognition test 3	3072	1432	1005	435	129	71

Source: authors' compilation from CNLSY79 and NLSY79.

Appendix Table 2 Statutory school entry ages in mathematics test score analysis sample

A. Number and frequency of sample members treated by each statutory age of school entry

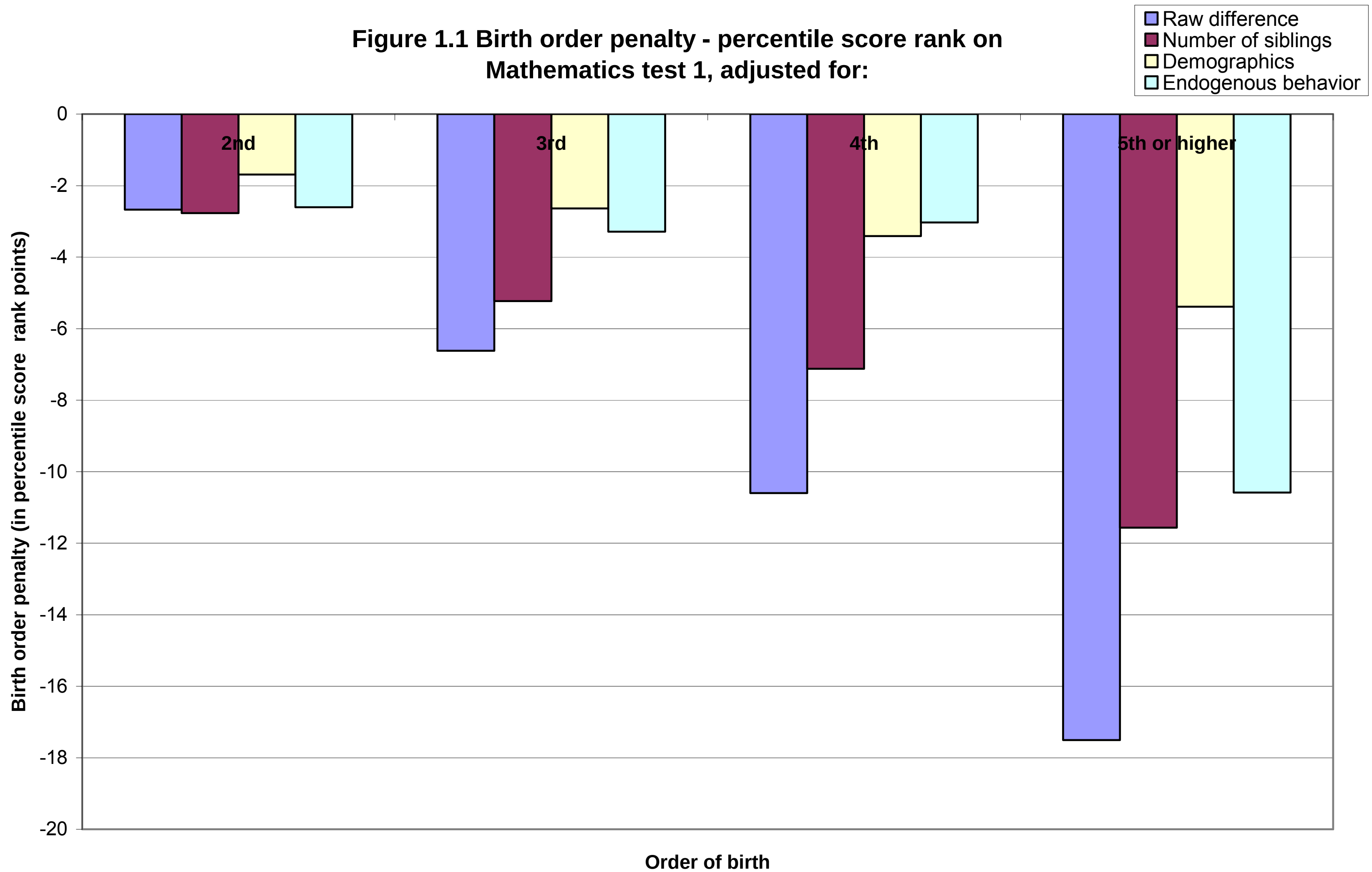
	Permitted age		Compulsory age	
Age range	Frequency	Percent	Frequency	Percent
3.83<= Entry age <4.00	618	10.98%	0	0.00%
4.00<= Entry age <4.25	1	0.02%	0	0.00%
4.25<= Entry age <4.50	40	0.71%	0	0.00%
4.50<= Entry age <4.75	33	0.59%	7	0.12%
4.75<= Entry age <5.00	1823	32.39%	328	5.83%
5.00<= Entry age <5.25	269	4.78%	44	0.78%
5.25<= Entry age <5.50	0	0.00%	0	0.00%
5.50<= Entry age <5.75	708	12.58%	722	12.83%
5.75<= Entry age <6.00	911	16.18%	1290	22.92%
6.00<= Entry age <6.25	817	14.51%	703	12.49%
6.25<= Entry age <6.50	0	0.00%	31	0.55%
6.50<= Entry age <6.75	0	0.00%	239	4.25%
6.75<= Entry age <7.00	8	0.14%	1231	21.87%
7.00<= Entry age <7.25	314	5.58%	654	11.62%
7.25<= Entry age <7.50	0	0.00%	0	0.00%
7.50<= Entry age <7.75	0	0.00%	15	0.27%
7.75<= Entry age <8.00	0	0.00%	278	4.94%
8.00<= Entry age	87	1.55%	87	1.55%

B. Summary statistics for statutory age of school entry in mathematics test score analysis sample

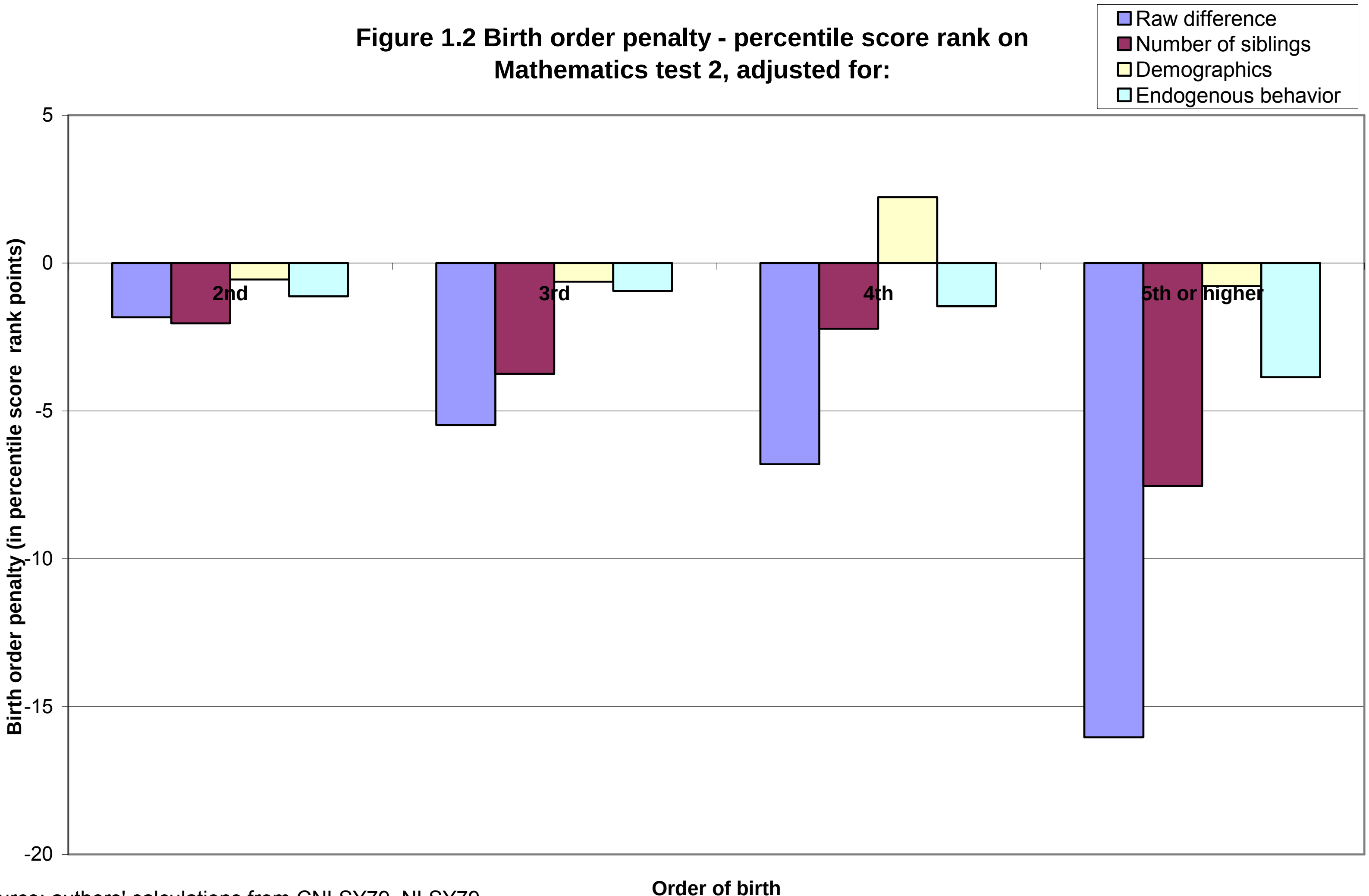
Statistic	Permitted age	Compulsory age
Mean	5.42	6.37
Median	5.67	6.00
Mode	5.00	6.00
Standard Deviation	0.83	0.75
Variance	0.68	0.56
Range	4.34	3.34
Interquartile Range	1.08	1.08
Number of respondents	5629	

Source: authors' calculations from Children of NLSY79 and Compulsory Schooling Law (CoSLaw) data base (compiled by authors).

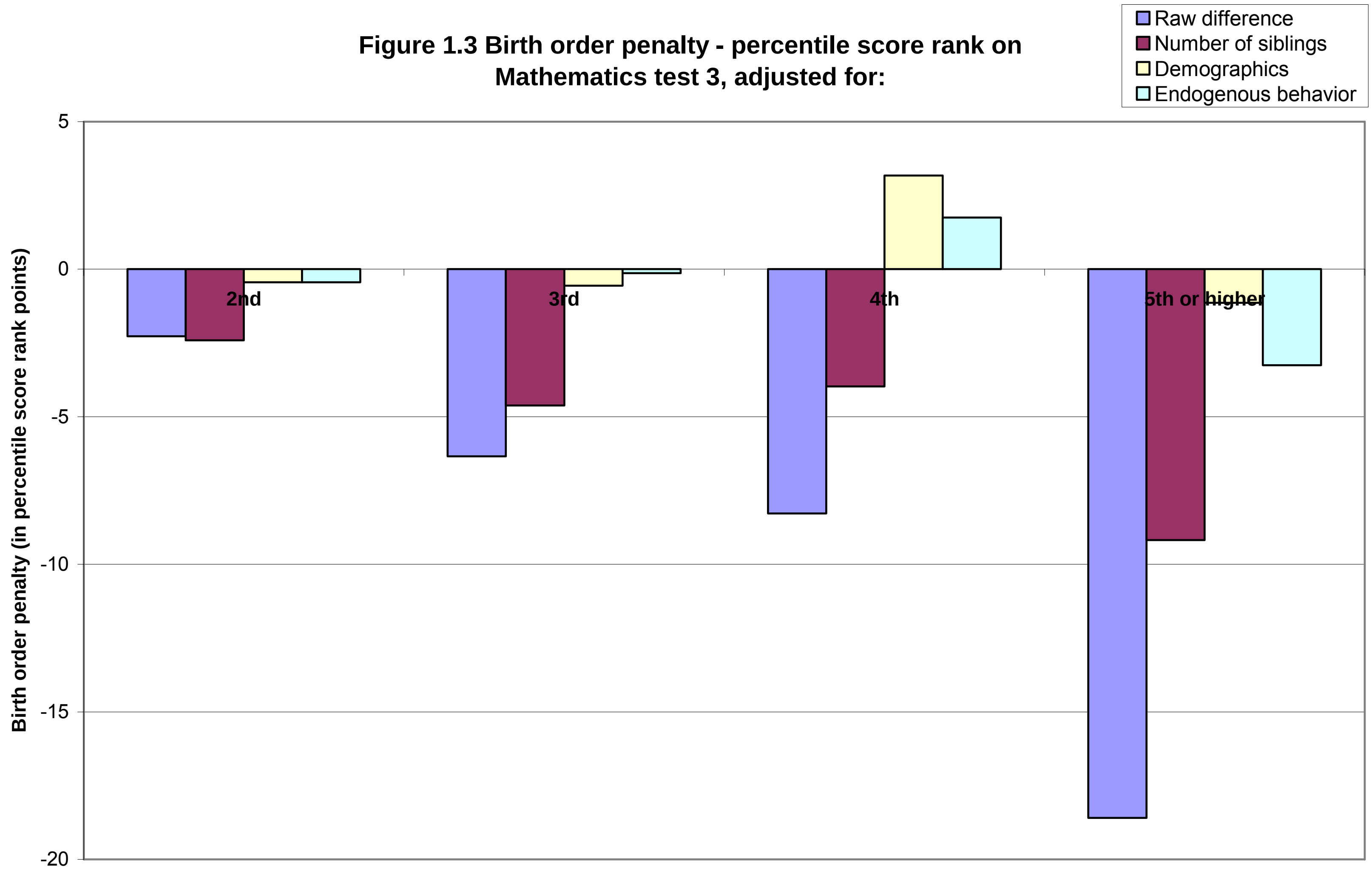
Figure 1.1 Birth order penalty - percentile score rank on Mathematics test 1, adjusted for:



Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 4

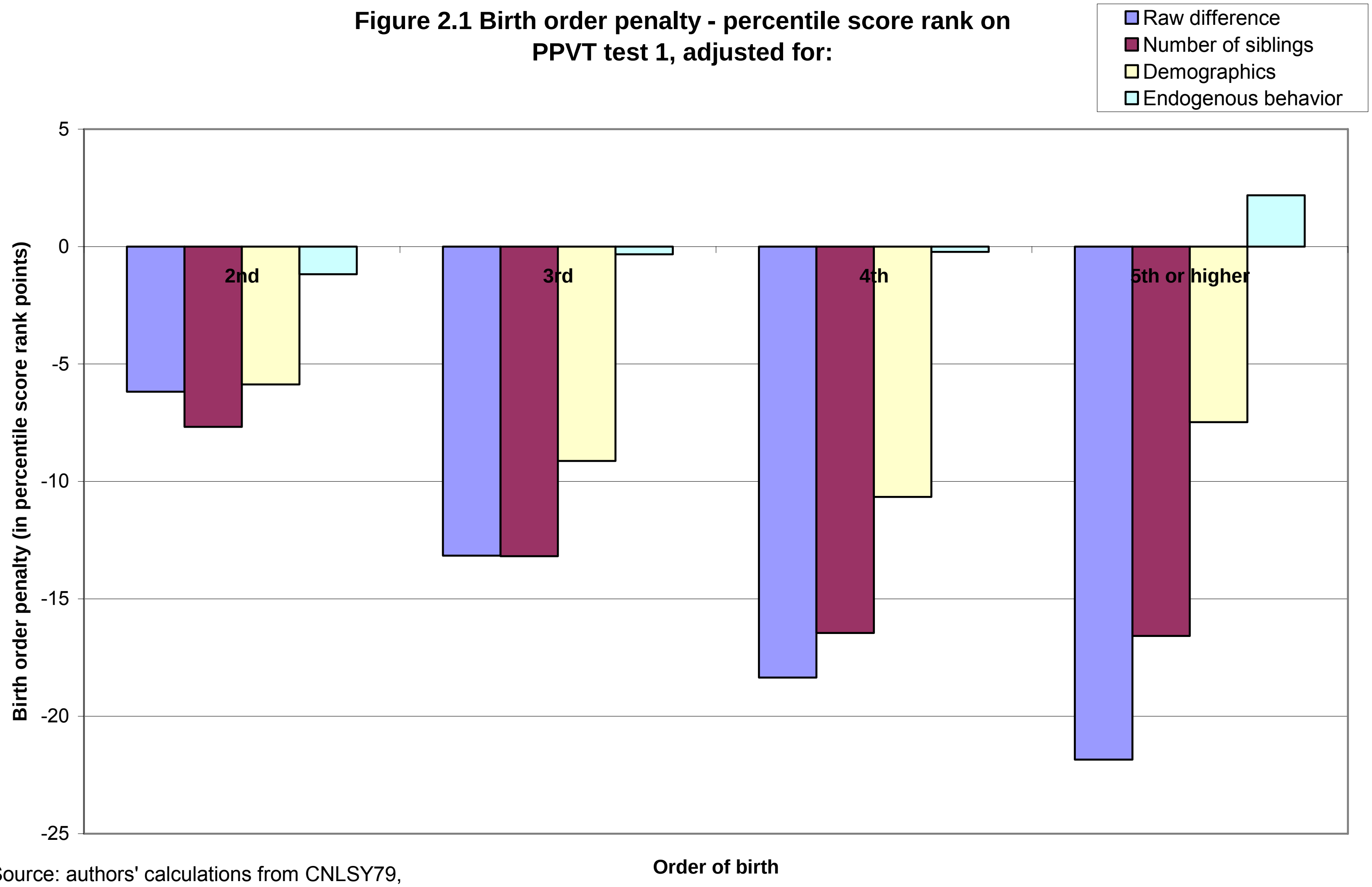


Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 4

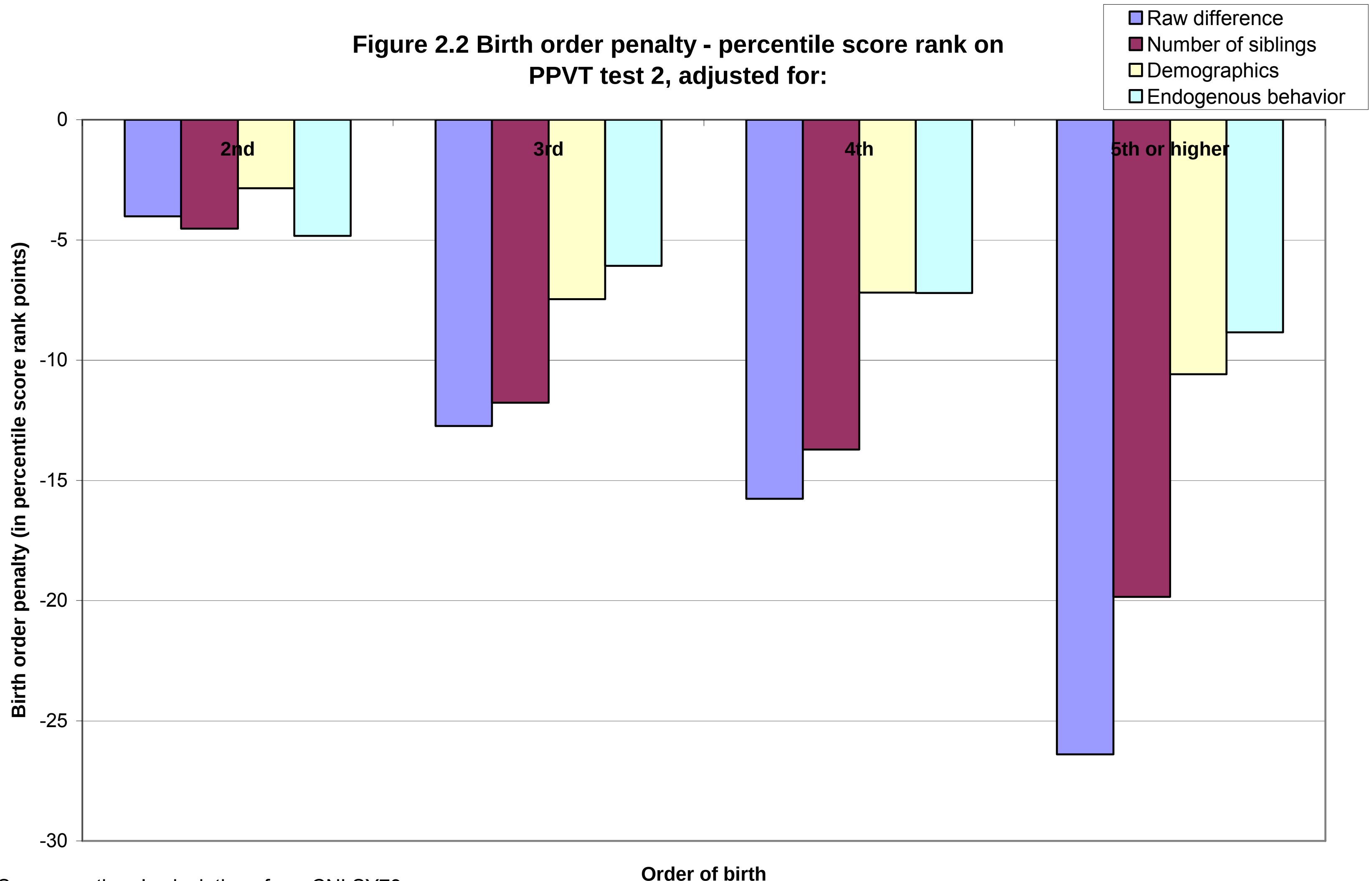


Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 4

Figure 2.1 Birth order penalty - percentile score rank on PPVT test 1, adjusted for:

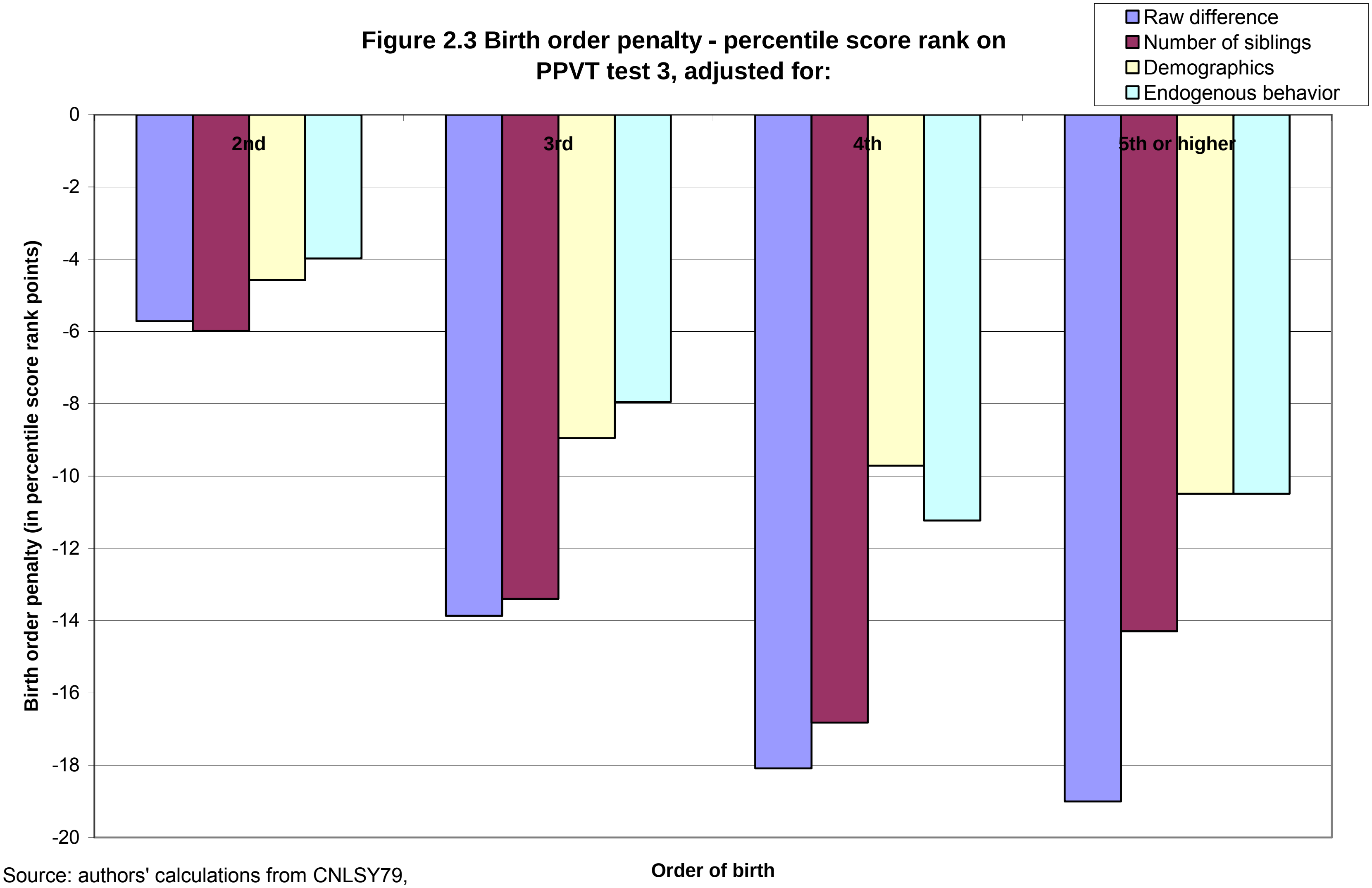


Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 5

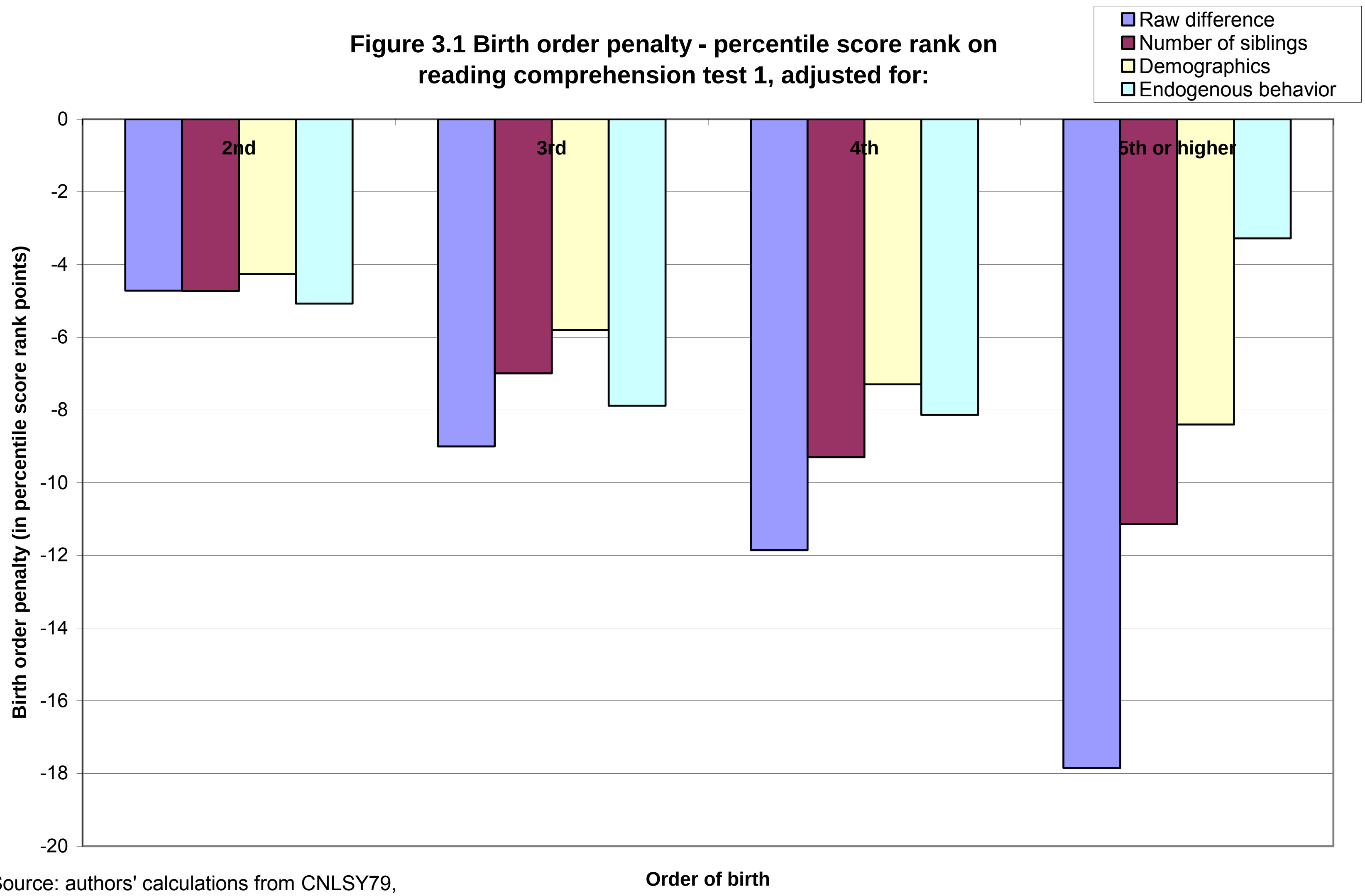


Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 5

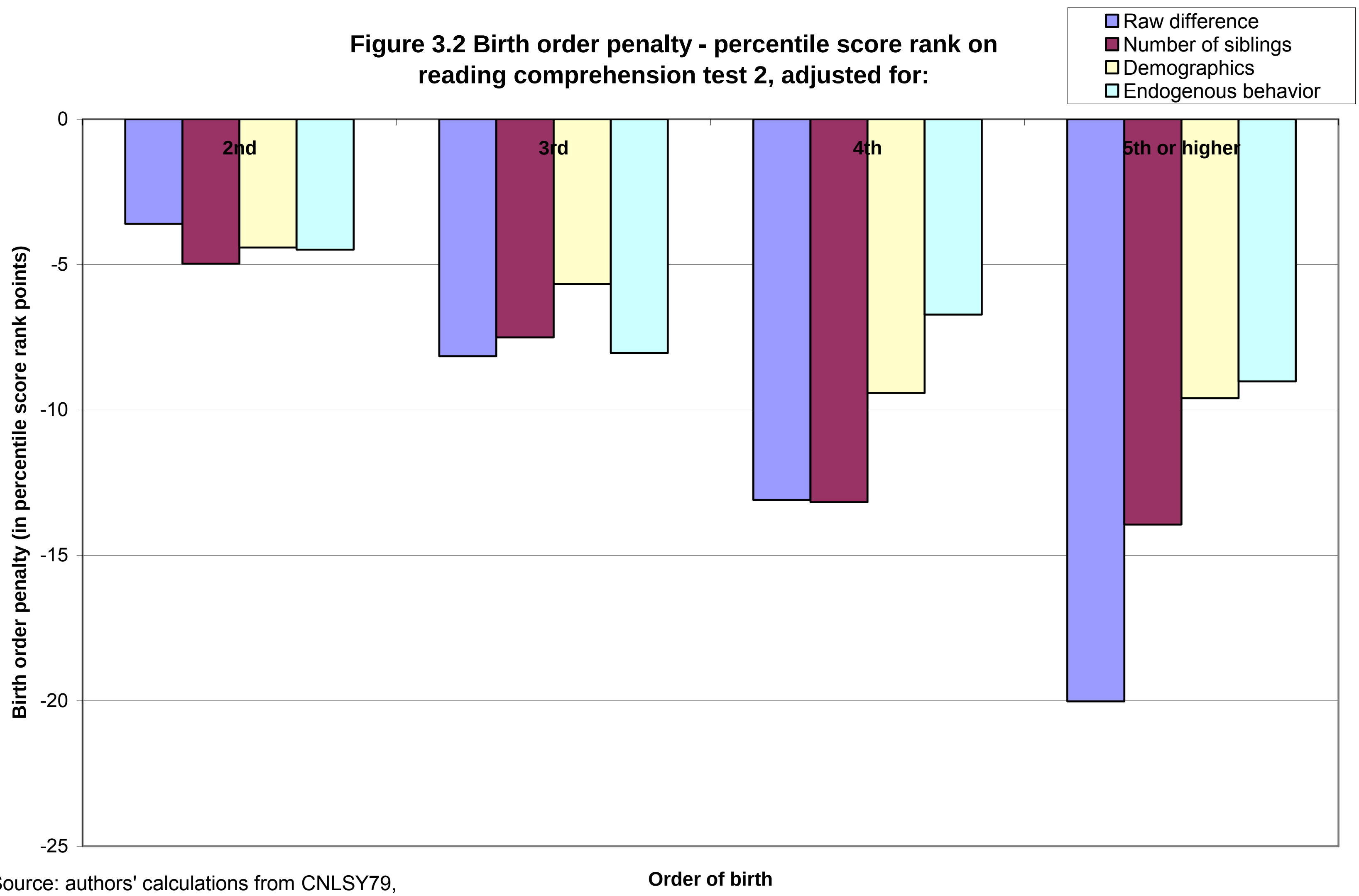
Figure 2.3 Birth order penalty - percentile score rank on PPVT test 3, adjusted for:



Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 5

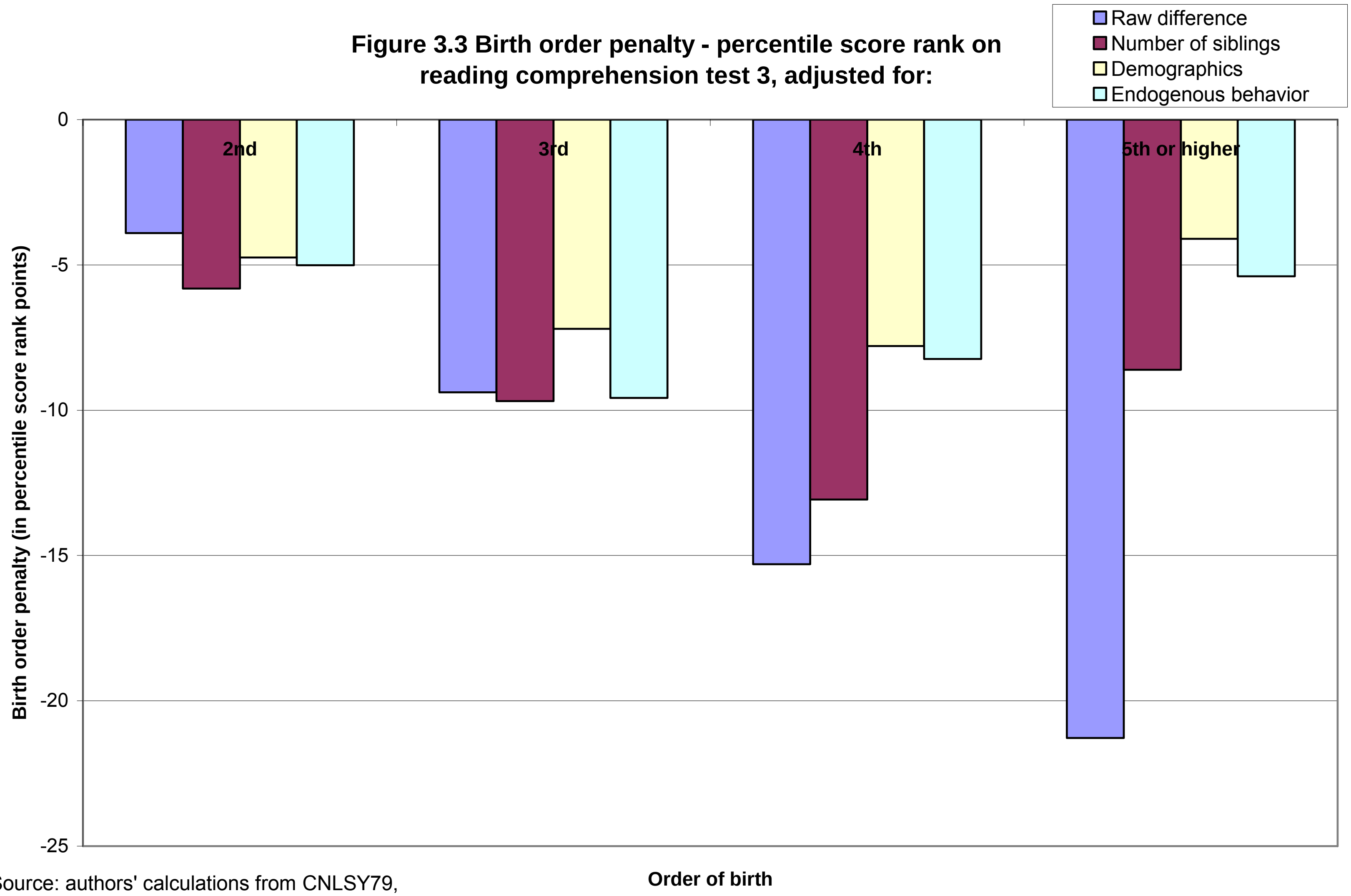


Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 6



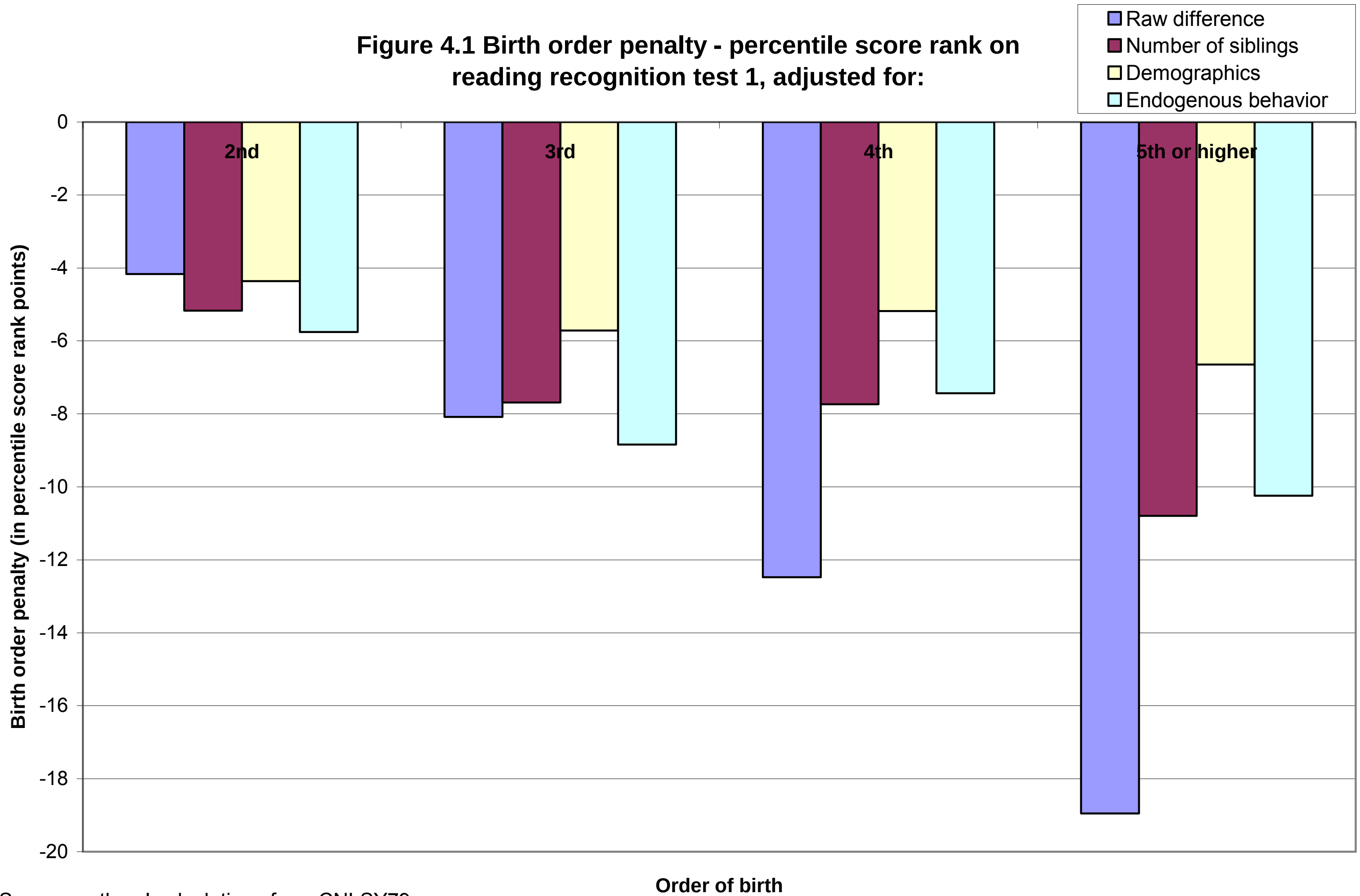
Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 6

Figure 3.3 Birth order penalty - percentile score rank on reading comprehension test 3, adjusted for:



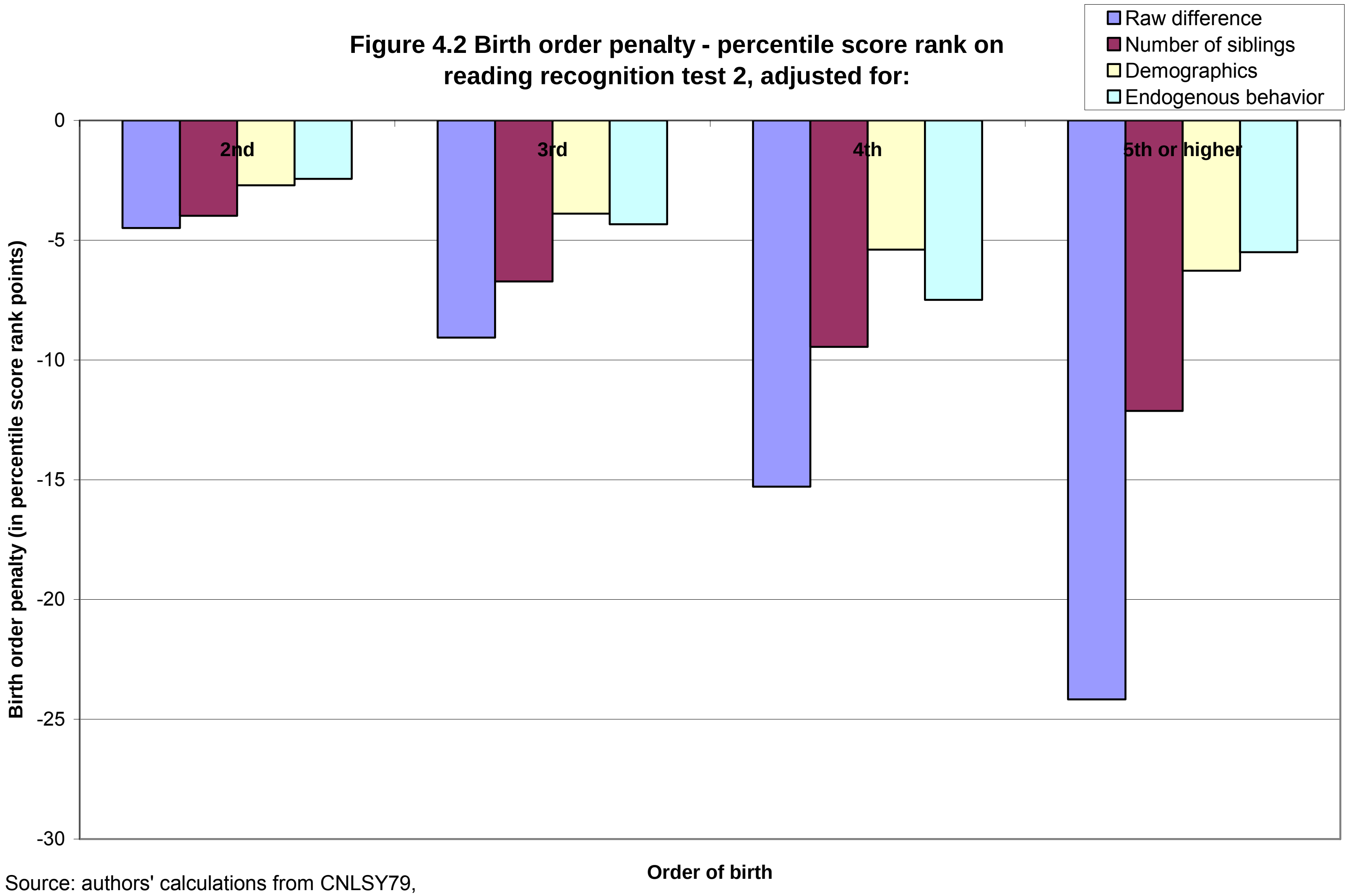
Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 6

Figure 4.1 Birth order penalty - percentile score rank on reading recognition test 1, adjusted for:

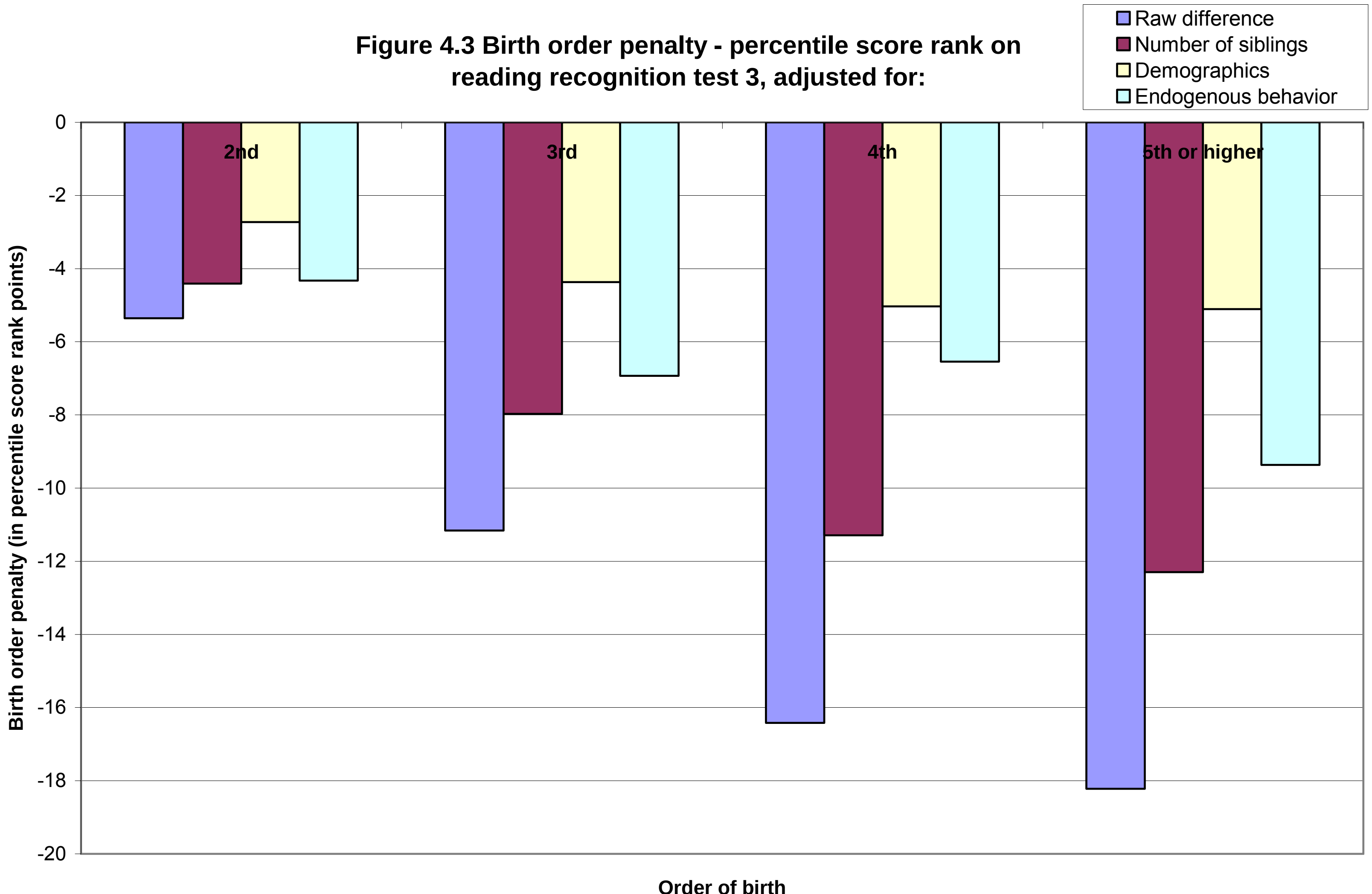


Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 7

Figure 4.2 Birth order penalty - percentile score rank on reading recognition test 2, adjusted for:



Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 7



Source: authors' calculations from CNLSY79, NLSY79, Table 1 and Table 7