Motherhood Delay and the Human Capital of the Next Generation

Amalia R. Miller*

October 24, 2008

Abstract

This paper exploits biological fertility shocks as instrumental variables to estimate the effect of motherhood delay on the cognitive ability of the next generation. Using detailed panel data on women in the NLSY79 and their first-born children aged 5 to 14, we find a year of delay leads to significant increases in math and reading scores: a 7 year delay produces gains on par with the black-white score difference. These results reveal a potential weakness of pro-natalist policies promoting early motherhood. While such policies may increase total period fertility rates, they will be less effective at increasing total human capital.

*Economics Department, University of Virginia, P.O. Box 400182, Charlottesville, VA, 22904. armiller@virginia.edu.

1 Introduction

Policymakers and scholars have expressed concern over recent increases in motherhood delay, primarily because of its impact on total period fertility rates, but also because of its potential harm to child and maternal health. Many pro-natalist policies implemented in Europe and elsewhere have produced the intended or unintended consequence of encouraging young women to begin childbearing sooner.[1] While such policies may increase the size of the next generation, their impact on that generation's total human capital are unclear, since per person human capital may change as well. In the spirit of Becker and Lewis (1973) and Hanushek (1992), this paper explores a potential national policy tradeoff, embodied in motherhood timing, between the quantity and quality of children. Using biological fertility shocks as instrumental variables, the causal effect of motherhood delay on the cognitive ability of first-born children is shown to be positive and significant.

Related tradeoffs, embodied in the motherhood timing decision, have been demonstrated at the individual level. Women who delay motherhood tend to achieve greater career success Miller (2006), but they also experience fewer years of motherhood and may face increased health risks for themselves and their children (Royer (2005), Alonzo (2002)). Furthermore, delaying motherhood beyond the years of peak fecundity can lead to difficulties in conception and reduced opportunity for childbearing.[2] This paper examines a new factor in the calculus of optimal motherhood timing: the effects of delay on the human capital of children.

This paper also expands our understanding of the production function for early human capital formation. A common challenge in this literature is identification, or determining the causal effects of the various home and school inputs that may influence cognitive devel-

[1] See Grant, Hoorens, Sivadasan, van het Loo, DaVanzo, Hale, Gibson, and Butz (2004) for extensive discussion of fertility incentives in Europe. In the United States, federal child tax benefits have increased in recent years and amount to over $140 billion annually (Mumford (2007)).

[2] Hewlett (2002) describes some personal costs of motherhood delay, and argues that many successful career women have come to regret the family sacrifices they made to further their professional goals.

opment.[3] This paper contributes to the literature by using a novel identification strategy to uncover the effects of a changing maternal characteristic. Unlike race, ethnicity, and innate genetic ability, maternal age at first birth reflects a choice on the part of women that may respond to financial incentives, cultural norms and policy environment. Three biological fertility shocks are employed as instrumental variables for motherhood timing: (1) whether first pregnancy ended in miscarriage, (2) whether conception of the first child occurred while using contraception, and (3) elapsed time from first conception attempt to first birth. They are plausibly exogenous factors outside of a woman's control that shift actual motherhood timing away from optimal or expected timing by about 6 months to 1.5 years.[4] The evidence from this approach builds on previous studies using cross-sectional or within-family comparisons, and expands the range of maternal ages considered beyond the teenage years. Furthermore, the paper provides evidence that the channel for the positive effect is not limited to changes in family composition or to higher income.

The key relationship studied in this paper is between motherhood timing and child cognitive ability, as measured by Peabody Individual Achievement Test (PIAT) mathematics and reading scores.[5] At the individual level, test scores have long been a strong predictor of educational attainment and earnings, and their importance has only increased in recent decades (Murnane, Willet, and Levy (1995), Neal and Johnson (1996), Currie and Thomas (2001)). At the national level, test performance in mathematics and science is strongly related to economic growth, and may matter more than years of completed schooling (Hanushek and

[3]Murnane, Maynard, and Ohls (1981) argues that the general dearth of suitable instruments available to researchers studying families is the result of a lack of compelling theory and the fact that many potential instruments are themselves choice variables.

[4]This identification strategy, similar in spirit to Hotz, Mullin, and Sanders (1997), was first employed in Miller (2006) to estimate the impact of motherhood timing on career path. Sections 4 and 5 of this paper provide evidence supporting the validity of the instruments in the current context of child cognitive outcomes.

[5]The PIAT is a well-known test that has been used extensively in studies of child cognitive development and of its impact on later outcomes. Test takers are not rewarded based on their performance. As a result, scores most likely reflect some combination of ability and motivation, or cognitive and non-cognitive components of human capital (Segal (2006)).

Kimko (2000), Barro (2001)).

Using data on first-born children aged 5 to 14 in the NLSY79 Children sample, this paper finds delayed motherhood leads to significant increases in PIAT test scores. The relationship is robust to the inclusion of various controls for observable elements of maternal human capital, as well as the use of instrumental variables to address the potential endogeneity of motherhood timing.

2 Background on Child Cognitive Development

This study contributes to the larger interdisciplinary literature on the determinants of child cognitive development. Although many previous studies have linked maternal characteristics and behaviors to child test scores, this is the first to establish the causal connection between motherhood timing and child ability using biological fertility shocks as instrumental variables. In addition, the paper contributes by exploring the potential channels for the effect.

Previous research into the relationship, consisting primarily of cross-sectional comparisons, finds test score benefits associated with motherhood delay.[6] In order to interpret the estimates as causal, however, it is necessary to assume that motherhood timing is exogenous, conditional on covariates such as education, test scores and family background. This assumption is invalid if women who are more devoted mothers choose to have children sooner or if women who are more ambitious and accomplished choose to delay. In the first case, prior estimates will understate the benefit from delay, while in the second case, they will overstate it. OLS will also be biased for older children if there is a correlation between unmeasured school quality and motherhood delay.

[6]See Roosa, Fitzgerald, and Carlson (1982). Other studies of test scores, not focused on the effects of motherhood timing, still include it as a control. Using data from the Early Childhood Longitudinal Study, Fryer and Levitt (2004) finds significant gains from delay in a model with an exceptionally wide range of covariates.

More recent studies exploit the sampling of the NLSY79 to conduct within family comparisons of children born to sisters who became mothers at different ages. Geronimus, Korenman, and Hillemeier (1994) associates teen motherhood with lower test scores using OLS models, but with higher scores in models with family fixed effects. However, the paper uses data only through 1990, and the result appears to be an artifact of the limited range of maternal ages represented in the early years of the data. When the dataset is extended beyond the first 3 waves of the Children survey, the family fixed effects framework yields positive effects of motherhood delay for first-born children that are statistically significant for reading, and robust to including controls.[7]

The main disadvantages of the within family comparison strategy are that a segment of the sample is lost and a great deal of important variation in the data is absorbed in the fixed effects, thus reducing the precision of the estimated effects of motherhood timing. At the same time, family fixed effects may not solve the endogeneity problem.[8] The validity of the estimates still relies on the assumption that differences in behavior between siblings are random. In this context, it means that a woman who begins childbearing younger is no different from her sister who delays in any unmeasured dimension that matters for child cognitive development, including parenting quality of either mother or father. If women who are more family oriented choose to begin childbearing before their sisters and also to invest more time and energy in playing with and educating their children, the bias will be negative, and benefits of delay understated. If more capable mothers, conditional on observables, delay

[7]See Table 4 in Turley (2003): in the sample of first-born cousins, the estimated effect of maternal age at first birth on child reading score drops from 0.078 (significant at 0.1%) in OLS to 0.054 (significant at 5%) with family fixed effects. Effects on math and vocabulary scores drop from 0.054 to 0.020 and 0.046 to 0.024, respectively. The full specification includes marital status and income as control variables. Since these variables are likely affected by motherhood timing, and may constitute an important channel for the effects, their inclusion as controls may bias the results against finding a beneficial effect of delay. However, in a pattern consistent with the the investigation of channels in Section 7 of this paper, the additional controls have little effect on the main estimates.

[8]This point is clearly acknowledged in the prior literature. See for example the discussions in Geronimus and Korenman (1992) and Rosenzweig and Wolpin (1995).

childbearing relative to their sisters, the bias will be positive. Sibling fixed effects models will also be biased if researchers fail to account for birth order effects, in this case, of the mothers (Black, Devereux, and Salvanes (2005)).

Since no single approach is ideal, and previous work has yielded mixed results, it is important for researchers interested in uncovering the effects of delay to accumulate evidence using various empirical strategies. This paper builds on the previous approaches by incorporating the possibility that motherhood delay is a choice that may be related to unmeasured individual-specific traits and life experiences that differ even within families.[9] Similar instrumental variables strategies using biological variation from miscarriages (Hotz, Mullin, and Sanders (1997)) and age at menarche (Klepinger, Lundberg, and Plotnick (1999)) as well as regional policy variation in welfare generosity and access to family planning services (also in Klepinger, Lundberg, and Plotnick (1999)) have yielded important insights into the effects of teenage motherhood on women's human capital and career outcomes. The biological shocks used as instrumental variables in this paper are selected for their potential to shift motherhood timing for women in their twenties and early thirties.

3 Empirical Framework

Test performance is generally modeled as the outcome of innate ability, family inputs and school inputs, cumulative to the time of performance measurement (Hanushek (2002), Todd and Wolpin (2003)). Parents invest in their children's development with time for instruction, play and monitoring, and money for market goods and services such as books, tutors, nutritious food, outside child-care and quality schooling. The benefits from motherhood delay can operate through various channels: (1) improved financial status; (2) changes in family

[9] If the instruments are valid, family fixed effects should be unnecessary. At the same time, it is theoretically possible to add fixed effects to instrumental variables model. Unfortunately, the results from this strategy are too imprecise to be informative: the mother age at first birth coefficient is 0.023 with a standard error of 0.633.

structure leading to increased quantity of parental resources per child; and (3) improved quality of parental inputs invested in child development.

The financial status of a family may improve with motherhood delay as a result of greater maternal earnings. Miller (2006) estimates that each year of motherhood delay during a woman's twenties leads to a 10 percent increase in her career earnings. The effective quantity of parental inputs may increase because of lower fertility and fewer younger siblings for the child, or improved family stability. Older adults, who have had more time to accumulate knowledge, experience and emotional maturity, may provide higher quality parental inputs.[10] Furthermore, older first-time mothers may be more perceptive of their children's idiosyncratic needs and more capable of managing those needs (as Murnane, Maynard, and Ohls (1981) hypothesizes for more educated mothers). On the other hand, motherhood delay may produce a reduction in maternal time inputs: Miller (2006) shows that later mothers return to paid work, and to full-time employment, sooner than earlier mothers.[11] Later mothers may also suffer from diminished energy and attention available for child-rearing.

The baseline empirical model of child test scores takes the following form:

$$Score_{it} = \beta_0 + \beta_{A1B} * MotherA1B_i + \beta_X * X_{it} + \epsilon_{it}$$

where i provides a unique index for each mother and oldest child pair, $MotherA1B_i$ is the

[10]Miller and Zhang (2007) illustrates the importance of both quality and quantity dimensions of maternal investment. In-depth observational studies comparing the parenting styles of mothers of different ages have produced somewhat mixed results. There is evidence of a positive association of age at first birth with "supportive maternal behaviors" such as praise and affection and a negative association with "aversive interactions" such as criticism and physical punishment (Conger, McCarty, Yang, Lahey, and Burgess (1984)). Barratt and Roach (1995) rated teen mothers in their sample as less appropriate in their interactions with their first-born infants and less vocally responsive than older mothers. However, Philliber and Graham (1981) detected no significant independent relationship between maternal age on various dimensions of interactions with children.

[11]Berger, Hill, and Waldfogel (2005) use NLSY79 data to show that women who returned to paid employment within a few months of childbirth were less likely to breastfeed or to provide their babies with preventative medical care. In OLS and Probit models on my sample, breastfeeding is positively associated with maternal age at first birth, but the relationship is negligible, negative and statistically insignificant under instrumental variables.

mother's age at first birth, and X_{it} includes child age at testing (flexibly captured by a series of indicator variables), sex, race and ethnicity. In the full regression model, the X_{it} matrix includes controls for maternal characteristics such as test scores, educational attainment, and substance use during pregnancy. The model is first estimated using ordinary-least-squares (OLS). Later, the exogeneity assumption on $MotherA1B_i$ is relaxed, and the model is estimated using the 3 biological fertility shocks as instrumental variables (IV) for motherhood timing.

In addition to establishing the overall relationship, this paper explores the primary channels for the effect. This is accomplished by estimating expanded versions of the empirical model that separately include measures of family income and family structure. Although these variables are themselves likely endogenous, it is useful to see how their inclusion influences the estimated impact of motherhood delay, β_{A1B}.

4 Data Description

Data are from two U.S. Bureau of Labor Statistics surveys: the National Longitudinal Survey of Youth 1979 (NLSY79), and the NLSY79 Children and Young Adults survey. The first survey contains a national sample of men and women aged 14-22 in 1979, at their first interviews. Respondents were re-interviewed annually until 1994, and then biennially. The survey contains an unusually rich set of information, including detailed labor market data and detailed pregnancy, childbirth and contraceptive use histories. The key variables taken from this survey are maternal characteristics, including age at first birth, AFQT score, educational attainment, biological fertility shocks and substance use during pregnancy, as well as the three instrumental variables.

The first instrument is an indicator for the woman's first non-aborted pregnancy ending in

miscarriage or stillbirth.[12] Medical evidence does not support a strong impact of behavioral factors on miscarriage risk. Rather, over 85% of miscarriages occur within the first trimester of pregnancy and over 90% are caused by genetic defects or other anomalies that prevent the fetus from developing properly (Porter (1999)). Since miscarriage has been associated with some extreme behaviors such as heavy alcohol use or drug addiction, control variables are added for substance use during first (or earliest reported) pregnancy, such as smoking cigarettes, drinking alcohol, and using marijuana or cocaine.

The second instrument, an indicator variable for contraceptive use at the time of pregnancy, is taken from a combination of two questions asked after each live birth. The first asks if the woman used contraception prior to the pregnancy, and the second, if she ceased contraception prior to conception. The indicator is intended to identify women who first became pregnant "accidentally" and in spite of precautionary efforts.

The third instrument, time to first conception, is constructed using biennial information on contraceptive use. Among women who report prior contraceptive use, a woman's first conception attempt is defined to start in the first year that she reports sexual activity and no contraceptive use. The lag is defined as the number of years between the woman's first attempt and the birth of her first child. For women who never report using contraception or whose first attempt is identified after their first birth, the variable is set to zero, making it a non-linear function of inferred conception attempt and motherhood timing. A companion dummy variable for "reported contraception prior to first birth" is included to remove the potential bias from contraceptive use.

Since the key variables are based on self-reported miscarriage and contraceptive use, there is a danger that systematic misreporting could compromise the identification strategy.

[12]Hotz, Mullin, and Sanders (1997) estimates bounds on the effects of teen motherhood that incorporates potential contamination of the miscarriage natural experiment from misreporting, and from latent abortion-types who miscarry. These bounds confirm the qualitative instrumental variables findings. The problem of latent abortion-types in the miscarriage group should be smaller in this paper, which focuses on older mothers, as abortion rates are substantially lower for later pregnancies.

Hence, it is worthwhile to confirm their validity with outside sources. Although one may be concerned about under-reporting of miscarriages, the 13% rate in the sample is in line with medical estimates of about 15% of all recognized pregnancies ending in miscarriage (Regan (2001)). While a direct comparison of contraceptive use variables is not available, some general features of the NLSY match data from independent sources such as the Alan Guttmacher Institute and the National Surveys of Family Growth (NSFG). The most popular type of contraception for childless women is the birth control pill, although the pill has been somewhat replaced by condoms for women in more recent cohorts.

The unintended pregnancy variable has a mean of 0.33 in the sample. This may seem high when compared with failure rates of properly used contraceptive methods, but is lower than the NSFG reported rates of unintended pregnancy (57% of pregnancies in 1987). The NLSY79 also asks if pregnancies were "desired" and women in the sample commonly report that their children were not planned. In fact, the fraction of children who are "intended" is quite low. For example, consider responses to the 1982 question, "Just before you became pregnant the first time, did you want to become pregnant when you did?" 11.6% of respondents said "Yes," 4.2% said "Didn't matter," 63.8% said "No, not at that time," and 20.5% said "No, not at all." These comparisons provide some corroboration for the accuracy of reported miscarriage and contraceptive use in the NLSY79, and support the claim that biological shocks play a role in human fertility.

Miller (2006) also shows that the unintended pregnancy instrument is not operating through differences in the type of contraception used. However, the consistency with which a woman used contraception could be a source of bias. Consistency may be correlated with the intensity of preferences against childbearing or with the organizational skills of the woman. The main regressions controls for mother's AFQT score and education, which together proxy for organizational skill. Furthermore, the main results are unchanged when additional controls are included for the mother's and her spouse/partner's expressed desire

for the child.[13]

Another concern that applies to all of the instruments is that the data may be contaminated by systematic misreporting of contraceptive use or miscarriage among women who intentionally aborted. Systematic misreporting, if present, would most likely be related to a woman's religious beliefs and attitude towards abortion, contraception and motherhood. A spurious correlation would exist between religion and the measured instrumental variables, while the true instruments would be uncorrelated. In fact, neither a woman's religious affiliation (at birth or in the present), nor her frequency of attendance at religious services, has any statistical power in predicting any of the three instruments. In regressions of each of the instruments on the basic controls and indicators for religion, the religion variables were statistically insignificant, both individually and jointly, which implies that misreporting is unsystematic in relation to beliefs.

The second survey is a biennial panel, starting in 1986, and interviewing the biological children of female NLSY79 respondents. This survey contains child demographic information such as sex, race and ethnicity, as well as the child cognitive outcome variables: Peabody Individual Achievement Test (PIAT) scores.[14] The test was administered in every survey year to children aged 5 and older. These scores are available separately for sub-tests that measure achievement in mathematics, reading recognition, and reading comprehension. The analysis makes use of age-specific standard scores. Average scores in the sample, shown in Table 1 are above the mean of 100 in the norming sample, and standard deviations are below the norming sample's value of 15. An overall scaled score is also computed for each test-taker by summing the scaled scores across the three subject areas. The total test score mean is

[13]The main regressions in Table 2 are repeated with 4 additional variables: 2 to indicate if the child was "desired" by his or her mother and by her spouse/partner, and 2 to indicate if the "desired" questions went unanswered. With the inclusion of these variables, the OLS and IV estimates remain positive and significant with values of 0.53 and 0.89, respectively.

[14]These are included as part of the NLSY79 Child survey. The PIAT is a widely-used assessment with demonstrated high reliability.

319.12 and standard deviation is 34.5.

The NLSY79 sample is restricted to women whose first biological child was interviewed in NLSY79 Children survey, and was born after 1982, the first survey year with contraceptive information. This restriction implies that test scores from 1986 are excluded from analysis. The analysis focuses on each woman's first-born child. This avoids the potentially confounding effects of birth order.[15] When considering later-born children, the current framework is unable to disentangle the impact of a mother's age at the time of her first birth from the impact of her age at the time of the tested child's birth and from the impact of fertility spacing. Spacing itself may be endogenous, as in Rosenzweig and Wolpin (1995), where later fertility decisions respond to observed birth outcomes from prior pregnancies. Hence, the results should be interpreted as applying to first-born children, and may not extend naturally to their younger siblings.[16]

The estimation sample is comprised of children between the ages of 5 and 14 at the time of PIAT assessment. Multiple test scores at different ages are available for most children: 140 are tested once, 219 are tested twice, 349 are tested three times, 532 are tested four times, and 174 are tested five times. To address the potential serial correlation in errors across repeated tests for the same child, standard errors are clustered at the child level in all regressions.[17] Summary statistics for the main variables in the estimation sample are presented in Table 1.

[15]Belmont and Marolla (1973) establishes the relationship between birth order and intelligence test scores for the cohort of Dutch men born between 1944 and 1947. The association does not depend on social class or spacing or fertility (Belmont, Stein, and Zybert (1978)). Kristensen and Bjerkedal (2007) argues that "social order" within a family is the true source of the effect.

[16]In a comparison of later-born cousins, Turley (2003) finds positive and significant effects of motherhood delay (measured by age at first birth). When the IV model in this paper is estimated on a sample including later-born children, the effect of Mother A1B is elevated to 1.01 (s.e. of 0.32). However, the interpretation of the coefficients is not as clear.

[17]In a robustness exercise reported in Table 5, we see that the main results are unchanged when estimation is conducted on a reduced sample of 521 children aged 5 or 6 at the time of testing.

5 Main Results

Before estimating the full model of the Section 3, it is useful to examine the raw relationships between test scores and maternal age in the data. Figure 1 plots the distributions of total test score for first-born children of women in each Mother A1B category: 18 to 20, 21 to 23, 24 to 26, 27 to 29, and 30 or older. Each figure shows a histogram of test score densities overlaid with a Kernel density curve. Two facts emerge. First, the distributions are largely overlapping. This implies that if one were to repeatedly draw children at random from each of the Mother A1B bins, the child with the youngest mother will frequently outperform the others. Motherhood timing does not appear to be the key determinant of test scores. Second, the distributions are noticeably different, and test scores for children born to older women have greater density at higher scores. The remainder of this section will focus on average differences in scores.

Table 2 reports the main regression results for first-born children. The effect of motherhood delay on test scores is positive and significant. The dependent variable is total standardized test score, and each column includes a set of indicators for child age (not reported for clarity). Tests were conducted biennially on children aged 5-14, yielding up to 5 scores for each child in the sample. Robust standard errors, clustered at the child level, are reported and used in hypothesis testing. The coefficient of interest, mother's age at first birth (A1B), decreases by a quarter when controls are added for child race and ethnicity, and by another two thirds when controls are added for observable maternal human capital, in the form of aptitude test score when young, and indicators for educational attainment. Controls for contraceptive use prior to and substance use during pregnancy are included in the next column to ensure to validity of the instrumental variables, but these have a negligible effect on the coefficient of interest. With the full set of controls, a year of motherhood delay is associated with a 0.533 increase in child total test score, statistically significant at the 1%

level.

If there are unobserved maternal traits that are related to both motherhood timing and to child ability, such as omitted aspects of maternal ability, ambition, and personality, the OLS estimates will be biased and inconsistent. We address this potential concern by instrumenting for Mother A1B with biological fertility shocks, largely unanticipated events, outside of the woman's control, that shift actual motherhood timing away from expected or optimal timing. The final column reports instrumental variables (IV) results, which are larger and less precisely estimated that OLS: a single year of motherhood delay leads to a 0.927 increase in test scores, significant at the 5% level. These magnitudes are non-trivial, equivalent to between 5% (OLS) and 10% (IV) of the test score difference between children of college graduates and those of high school dropouts. A 7 year delay produces gains on par with the black-white score difference. At the same time, as suggested by Figure 1, motherhood timing is clearly not the main source of test score differences. A standard deviation increase in Mother A1B will lead to 11% of a standard deviation increase in total test scores. In fact, the R-squared value implies that the entire model explains only about a quarter of the observed test score variation.

The instrumental variables have strong and reasonable estimated effects in the first stage regression predicting motherhood timing:

$$MotherA1B_i = \beta_0 + \beta_Z * Z_i + \beta_X * X_{it} + \epsilon_{it}$$

where i provides a unique index for each mother and oldest child pair, $MotherA1B_i$ is the mother's age at first birth, X_{it} is the vector of controls in the main regression equation, and Z_i is the vector of instrumental variables excluded from the main equation. Failed contraception reduces maternal A1B by 0.83 years (s.e. of 0.18, significant at .1 percent), duration of first observed conception attempt increases it by 0.79 (s.e. of 0.036, significant

at .1 percent), and miscarriage at first pregnancy increases it by 0.49 (s.e. of 0.25, significant at 5.3 percent). For a typical woman, the instruments cause about 6 months to 1.5 years of fertility shifts. These are small, but not trivial, effects that provide identification in the second stage, where the Mother A1B coefficient should be interpreted as the average effect of a year of motherhood delay due to biological fertility shocks. The instruments are jointly highly significant, with an F-statistic of 180.

Miller (2006) also provides evidence supporting the exogeneity of instruments. They are uncorrelated with a woman's religious affiliation or her attendance at religious services. While they predict wages for mothers, they fail to predict career outcomes for women prior to motherhood. This suggests that the IV estimates of career effects are not operating through a correlation between the instruments and fixed omitted maternal factors. As a further validity check in this paper, we exploit the fact that there are three instruments for a single endogenous variable, and test the implied over-identification restrictions. Hansen's J test statistic and its corresponding P-value are reported below each IV estimate. The null hypothesis is that the additional instruments are exogenous (under the assumption that at least one is valid), and the high P-values reported in the tables indicate failures to reject the null. Furthermore, IV estimation is repeated omitting one of the individual instruments in turn. The positive relationship is confirmed, although it is not always statistically significant.[18]

The next two tables show OLS and IV results from estimating the full model separately for each of the three test subjects: mathematics, reading recognition and reading comprehension. The OLS estimates in Table 3 are positive and significant for each subject separately, and smallest for reading comprehension. The IV estimates in Table 4 are positive for each subject, but not statistically significant for reading comprehension. Some variables, such as

[18]Omitting miscarriage leads to a point estimate for Mother A1B of 0.997 (s.e. of 0.41), omitting duration of first conception attempt leads to 0.623 (s.e. of 1.41), and omitting failed contraception leads to 0.911 (s.e. of 0.41). When the model is estimated on the larger sample of all children within the age range 5-14, born after 1982, the effect of Mother A1B is positive and remains statistically significant at conventional levels (1% to 8%) with the full set of instruments or after omitting any single IV.

mother's AFQT score and college graduation, have similar positive effects on scores in each of the three subjects. Others, such as child sex, vary across subjects; girls outperform boys in reading recognition, but lag in mathematics. Somewhat surprising are the large positive coefficients for cocaine use during pregnancy: since only 17 mothers in the sample report cocaine use, the variable may be unduly influenced by outliers.[19] These results show that the benefits from motherhood delay are present for both reading and mathematics.

6 Robustness Tests: Sampling and Weights

One may be concerned that the positive findings presented above are driven by some unusual property of the sample. Results from additional sample robustness exercises are reported in Table 5. First, the children of the NLSY79 women are not a representative sample of children in the United States. The first panel of the table presents estimates of the main coefficient of interest, under least-squares and IV, using observation weights provided by the Bureau of Labor Statistics. The weighted estimates are slightly smaller than their unweighted counterparts in columns 4 and 5 of Table 2, 0.444 instead of 0.533, and 0.900 instead of 0.927, but they remain economically and statistically significant at conventional levels.

A related concern is that children of older mothers may be at an increased risk for adverse health outcomes such as learning disabilities, and systematically less likely to take the PIAT. Trimming the left-tail of the score distribution for children of older mothers would artificially inflate their average scores. This seems unlikely in the data, however, since missing test scores (for a child in the sample) are not systematically related to Mother A1B: under Probit and IV-Probit models, the relationship is statistically insignificant (P-values of 0.48 and 0.877) and inconsistent in direction.

[19] Excluding these women from the estimation sample and dropping the control for cocaine use has no effect of the main results.

As an artifact of the sampling procedure, in which the children of a fixed cohort of women are surveyed, the older children in the sample are more likely born to women with lower A1B. To ensure that this feature is not introducing bias, the next two panels report results from estimates on a reduced sample of test scores: those achieved by children ages 5-8 at the time of testing, and by children ages 5-6. The last group includes only 1 score for each child, and includes children born between 1983 and 1999 to women between 18 and 40 years old. Again, the main results are confirmed, and somewhat strengthened on the smaller sample.

Finally, one may be concerned that an unusual group of women, either teenage mothers or women who delayed childbearing into their thirties, are the driving source of the findings. In the next two panels, this is shown not to be the case. When teen mothers are excluded, and when the sample is limited to women with A1B between 20 and 29, motherhood delay continues to improve children's test scores. The former restriction eliminates 149 observations, and highlights a key contribution of this paper: while previous studies have focused on the effects of teenage pregnancies, this paper considers the effects of delay for women at older ages. The latter restriction is particularly useful in light of the fact that female fecundity is known to decline with age. The drop is highly non-linear, and is most prominent after the mid-30s. Verifying the results on the limited range of Mother A1B provides some reassurance that the miscarriage and time to conception variables are not being affected by age at first pregnancy attempt.

7 Channels for the Main Effect

This section aims to uncover the underlying channels through which motherhood delay improves child human capital. In particular, we use observable controls to explore the first two channels listed in Section 3: improved financial status, and changes in family structure. This exercise is important for scientific and policy motivations. As an example of the latter,

consider the implications if the benefits to children from motherhood delay run exclusively through financial status. A policy that promotes early motherhood through a tax break or subsidy can also be designed to transfer sufficient wealth to the parents to partially or fully offset the loss in income. In that case, the next generation's total human capital will grow. If, instead, the primary channel is not financial, subsidies will be less effective (depending on the degree of substitutability) at protecting children's human capital.

The results in Table 6 provide evidence that income is not the primary channel. The first approach to income draws on a key finding from Miller (2006): wage gains from delay are concentrated among women with college degrees. By implication, if the test score benefit from motherhood delay flows from increased maternal lifetime income, we should expect to see the test score gains from motherhood delay also concentrated among college educated women. The new models augment the basic OLS model with an interaction term between A1B and the college graduate indicator, and the IV model with that interaction and similar interactions with each instrumental variable. The results in the first two columns show that A1B matters for all children, regardless of their mother's education level. Although the interaction terms have positive coefficients, these are not statistically significant, and the A1B coefficients are comparable in size to the original estimates in Table 2.

The second approach to income is to include an additional control for mother's wage rates in the year prior to the birth of the child. The new estimates are valid for inference on the more limited set of children whose mothers worked in the year before their birth. The restricted sample size is reduced from 4623 to 4019. If pre-motherhood wage rates are positively correlated with an unobserved element of maternal human capital that is transmitted to children, the coefficient on wages is biased upward. The sign of the bias is reversed under the scenario that women with higher ability spouses or partners earn lower wage rates, possibly due to less accumulated work experience. The estimated effect of pre-motherhood wage rates is statistically insignificant in both columns 3 and 4 of the table. The

point estimates are small: it would take an increase in wages of $7/hour (or a doubling of the average wage) to increase total test scores by 0.26 to 0.35. Inclusion of pre-motherhood wage rates leaves the main estimates on Mother A1B largely unchanged in magnitude or significance. Thus, the first two sets of estimation results in Table 6 provide evidence that maternal income is not the key channel for the effects of motherhood delay. In the final specification, the income of the father (spouse or partner of the mother) is included as a control, along with an indictor for missing income information. Again, the Mother A1B estimates are largely unchanged: the IV estimate is 0.924 (standard error of 0.41) and the OLS estimate is 0.478 (s.e. of 0.204).

Turning to the second proposed channel, Table 7 presents results from regressions that include controls for family structure. The first column reports OLS results with all three family structure controls: indicators for mother married before first birth and mother never divorced or widowed, and a count of the number of children born to the mother. Only family disruption has a significant relationship with test scores, with a 3.613 score advantage for children in stable households. The remaining columns contain IV estimates that echo the OLS findings. The additional controls are added in turn and then together, with only marital disruption showing an association. One should hesitate before interpreting these effects as causal, however, since prenatal risk factors and child cognitive development can themselves influence marital stability and later fertility. The key insight from the table is that the inclusion of these variables does little to alter the main coefficient estimate for A1B, which remains significant. The joint implication of this table and Table 6 is that the mechanism for the motherhood delay effect lies primarily in the unexamined channels and not in financial status or family structure. Although it may be tempting to attribute the residual effect to the third remaining channel, the quality of maternal inputs, obtaining direct evidence remains a goal for future research.

8 Conclusion

Although it is well known that motherhood delay has increased in the United States and in other countries in recent decades (Chen and Morgan (1991)), the implications of this demographic shift are not fully understood. This paper shows that motherhood delay leads to improved test scores for first-born children. This effect is present even after controlling for observable and unobservable dimensions of maternal human capital. Although financial status and family structure may be channels for this effect, they do not appear to be dominant. These results have direct implications for individual women considering their optimal timing of motherhood.

The results also suggest the presence of a "quantity-quality" tradeoff that may potentially reduce the effectiveness of pro-natalist policies aimed at increasing the human capital of the future generations by promoting early motherhood. Although the direction of the effect is clear, there is no direct way to extrapolate the values in this paper to predict the effects of policy changes. The estimates in this paper are based on a cross-sectional comparison of women in a given cohort. The biological fertility shocks are local in that their realizations independently shift motherhood timing for each individual woman. This is in contrast to state or national tax incentives or rewards for early childbearing that affect timing for many women at once. Peer effects in human capital acquisition can amplify the benefits from large-scale motherhood delay. At the same time, general equilibrium effects may reduce the benefits from delay by increasing demand and prices for factors of production in children's human capital, or by increasing supply of skilled labor and reducing the financial returns to ability. These factors may influence the scale of the national tradeoff between early motherhood and future human capital, but none will eliminate it altogether.

References

Alonzo, A. (2002). Long-Term Health Consequences of Delayed Childbirth. *Women's Health Issues 12*, 37–45.

Barratt, M. S. and M. A. Roach (1995, January-March). Early Interactive Processes: Parenting by Adolescent and Adult Single Mothers. *Infant Behavior and Development 18*(1), 97–107.

Barro, R. (2001, May). Human Capital and Growth. *American Economic Review 91*(2), 12–17.

Becker, G. and H. G. Lewis (1973, March - April). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy 81*(2), S279–S288.

Belmont, L. and F. Marolla (1973, December). Birth Order, Family Size, and Intelligence. *Science 182*(4117), 1096–1101.

Belmont, L., Z. Stein, and P. Zybert (1978, December). Child Spacing and Birth Order: Effect on Intellectual Ability in Two-Child Families. *Science 202*(4371), 995–996.

Berger, L. M., J. Hill, and J. Waldfogel (2005, February). Maternity Leave, Early Maternal Employment and Child Health and Development in the US. *The Economic Journal 115*, F29–F47.

Black, S., P. Devereux, and K. Salvanes (2005, May). The More the Merrier? The Effect of Family Size and Birth Order on Childrens Education. *Quarterly Journal of Economics*, 669–700.

Chen, R. and S. P. Morgan (1991). Recent Trends in the Timing of First Births in the United States. *Demography 4*, 513–533.

Conger, R., J. McCarty, R. Yang, B. Lahey, and R. Burgess (1984, May). Mother's Age as a Predictor of Observed Maternal Behavior in Three Independent Samples of Families.

Journal of Marriage and the Family 46(2), 411–424.

Currie, J. and D. Thomas (2001). Early Test Scores, Socioeconomic Status, School Quality and Future Outcomes. *Research in Labor Economics 20*, 103–132.

Fryer, R. and S. Levitt (2004, May). Understanding the Black-White Test Score Gap in the First Two Years of School. *Review of Economics and Statistics 86*(2), 447–464.

Geronimus, A. and S. Korenman (1992, November). The Socioeconomic Consequences of Teen Childbearing Reconsidered. *Quarterly Journal of Economics 107*(4), 1187–1214.

Geronimus, A., S. Korenman, and M. Hillemeier (1994, September). Does Young Maternal Age Adversely Affect Child Development? Evidence from Cousin Comparisons in the United States. *Population and Development Review 20*(3), 585–609.

Grant, J., S. Hoorens, S. Sivadasan, M. van het Loo, J. DaVanzo, L. Hale, S. Gibson, and W. Butz (2004). Low Fertility and Population Aging: Causes, Consequences and Policy Options. Technical report, RAND Corporation.

Hanushek, E. (1992, February). The Trade-off between Child Quantity and Quality. *Journal of Political Economy 100*(1), 84–117.

Hanushek, E. (2002). *Publicly Provided Education*. North-Holland.

Hanushek, E. and D. Kimko (2000, December). Schooling, Labor-Force Quality, and the Growth of Nations. *American Economic Review 90*(5), 1184–1208.

Hewlett, S. A. (2002). *Creating a Life: Professional Women and the Quest for Children*. Hyperion.

Hotz, J., C. Mullin, and S. Sanders (1997, October). Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analysing the Effects of Teenage Childbearing. *Review of Economic Studies 64*(4), 575–603.

Klepinger, D., S. Lundberg, and R. Plotnick (1999, Summer). How Does Adolescent Fertility Affect the Human Capital and Wages of Young Women? *Journal of Human Resources 34*(3), 421–448.

Kristensen, P. and T. Bjerkedal (2007, June). Explaining the Relation Between Birth Order and Intelligence. *Science 316*(5832), 1717.

Miller, A. (2006). The Effects of Motherhood Timing on Career Path. Mimeo, University of Virginia.

Miller, A. and L. Zhang (2007). The Effects of Welfare Reform on the Academic Performance of Children in Low-Income Households. Mimeo.

Mumford, K. (2007, February). The Optimal Tax Treatment of Families with Children.

Murnane, R., R. Maynard, and J. Ohls (1981, August). Home Resources and Children's Achievement. *Review of Economics and Statistics 63*(3), 369–377.

Murnane, R., J. Willet, and F. Levy (1995, May). The Growing Importance of Cognitive Skills in Wage Determination. *Review of Economics and Statistics 77*(2), 251–266.

Neal, D. and W. Johnson (1996, October). The Role of Premarket Factors in Black-White Wage Differences. *Journal of Political Economy 104*(5), 869–895.

Phillibert, S. and E. Graham (1981, February). The Impact of Age of Mother on Mother-Child Interaction Patterns. *Journal of Marriage and the Family 43*(1), 109–115.

Porter, R. (Ed.) (1999). *Abnormalities of Pregnancy* (17th ed.). Merck Research Laboratories.

Regan, L. (2001). *Miscarriage: What Every Woman Needs to Know*. Orion Books Ltd.

Roosa, M., H. Fitzgerald, and N. Carlson (1982, Spring). Teenage Parenting and Child Development: A Literature Review. *Infant Mental Health Journal 3*(1), 4–18.

Rosenzweig, M. and K. Wolpin (1995, March). Sisters, Siblings, and Mothers: The Effect of Teen-age Childbearing on Birth Outcomes in a Dynamic Family Context. *Econometrica 63*(2), 303–326.

Royer, H. (2005). Does Maternal Age Affect Infant Health?: Evidence from Sibling Comparisons. Mimeo, University of Michigan.

Segal, C. (2006). Motivation, Test Scores, and Economic Success. Mimeo, Harvard Business School.

Todd, P. and K. Wolpin (2003, February). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal 113*, F3–F33.

Turley, R. L. (2003, March/April). Are Children of Young Mothers Disadvantaged Because of Their Mother's Age or Family Background? *Child Development 74*(2), 465–474.

Figure 1: Distribution of Total Test Scores by Mother A1B

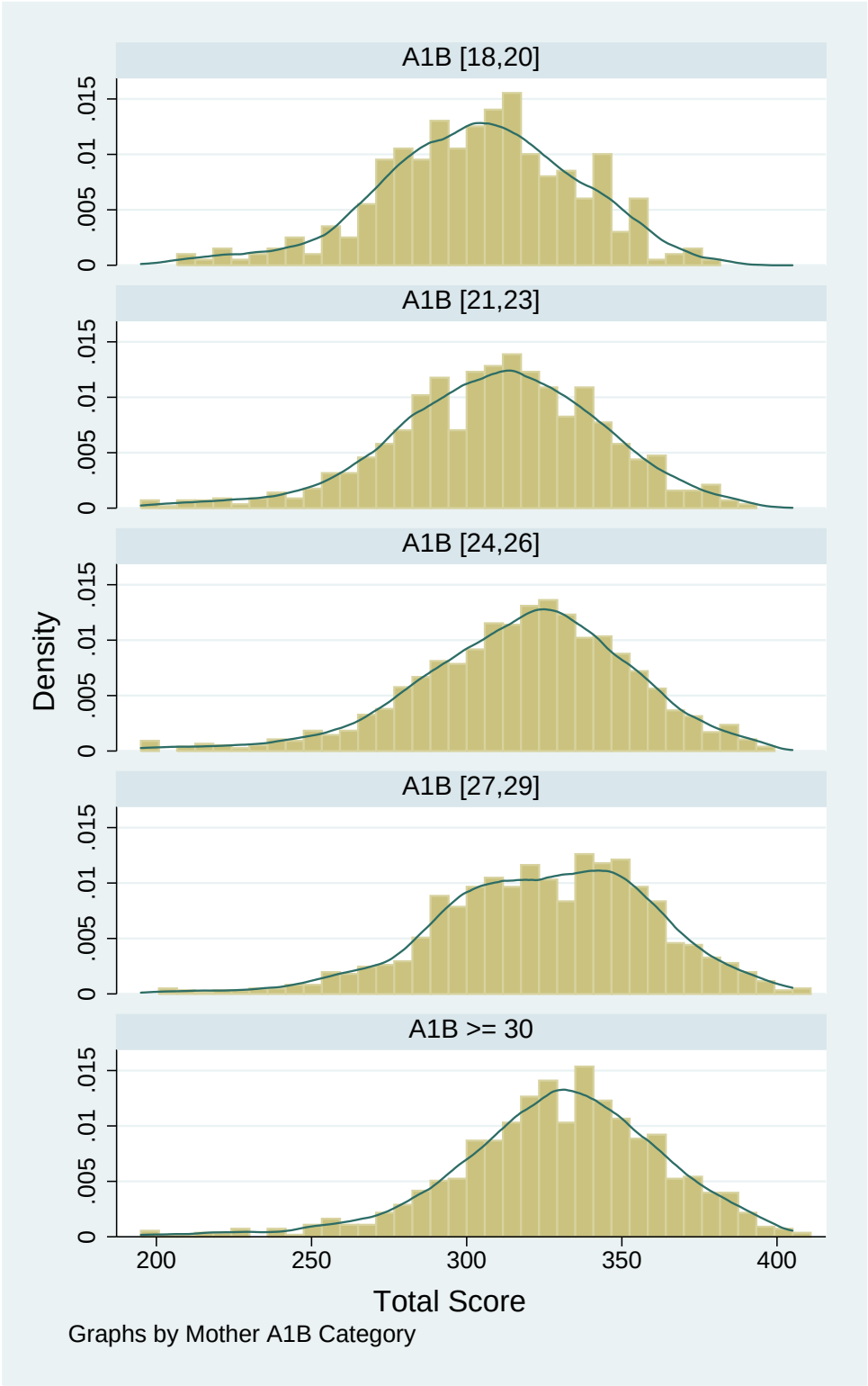

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	N
Mathematics score	105.150	13.681	4623
Reading recognition score	108.834	13.675	4623
Reading comprehension score	105.138	13.144	4623
Total test score	319.122	34.480	4623
Mother A1B	26.046	3.936	4623
Mother HS grad	0.532		4623
Mother college grad	0.409		4623
Miscarriage at first pregnancy	0.132		4623
Conceived using contraception	0.334		4623
Years to first conception	1.291		4623
Contraception	0.470		4623
Smoking during pregnancy	0.252		4623
Alcohol during pregnancy	0.520		4623
Marijuana during pregnancy	0.027		4623
Cocaine during pregnancy	0.012		4623
Child female	0.523		4623
Child Hispanic	0.173		4623
Child Black	0.217		4623

Source: NLSY79 and NLSY79 Children surveys. Sample is restricted to NLSY79 women and their first-born children, born after 1982, and belonging to the NLSY79 Children sample.

Table 2: Motherhood Timing and Total Test Scores

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV
Mother A1B	2.003***	1.568***	0.534***	0.533***	0.927**
	(0.20)	(0.20)	(0.20)	(0.20)	(0.41)
Child female		-1.408	-0.369	-0.325	-0.375
		(1.60)	(1.46)	(1.46)	(1.46)
Child Hispanic		-12.19***	-1.570	-2.078	-1.868
		(2.05)	(2.03)	(2.04)	(2.04)
Child Black		-19.74***	-7.299***	-7.037***	-6.876***
		(2.23)	(2.21)	(2.22)	(2.21)
Mother AFQT percentile			0.447***	0.434***	0.422***
			(0.038)	(0.039)	(0.040)
Mother HS grad			6.960**	6.699**	6.520**
			(3.24)	(3.25)	(3.24)
Mother college grad			10.55***	10.34***	9.617**
			(3.60)	(3.68)	(3.72)
Contraception				-0.542	-1.363
				(1.48)	(1.70)
Smoking during pregnancy				-0.866	-0.848
				(1.78)	(1.77)
Alcohol during pregnancy				2.496	2.399
				(1.54)	(1.55)
Marijuana during pregnancy				-2.224	-2.401
				(4.56)	(4.61)
Cocaine during pregnancy				15.34**	15.20**
				(6.02)	(6.00)
Constant	261.8***	280.8***	272.9***	272.6***	273.6***
	(5.73)	(6.41)	(6.67)	(6.75)	(10.5)
Observations	4623	4623	4623	4623	4623
R^2	0.08	0.14	0.26	0.26	0.26
Over-identification test of instruments					
Hansen's J statistic					1.504
P-value					0.471

Each regression includes the full set of age fixed effects. Coefficients suppressed for readability.
Robust standard errors, clustered at individual child level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: OLS Estimates of Motherhood Timing and Scores by Subject

Dependent Variable	(1) Mathematics	(2) Reading recog.	(3) Reading comp.
Mother A1B	0.212***	0.172**	0.149**
	(0.077)	(0.083)	(0.072)
Child female	-1.860***	1.107*	0.429
	(0.55)	(0.60)	(0.50)
Child Hispanic	-2.648***	0.467	0.104
	(0.81)	(0.85)	(0.70)
Child Black	-4.271***	-0.694	-2.072***
	(0.85)	(0.89)	(0.73)
Mother AFQT percentile	0.145***	0.147***	0.142***
	(0.015)	(0.015)	(0.013)
Mother HS grad	1.921*	2.399*	2.379**
	(1.14)	(1.40)	(1.10)
Mother college grad	3.390**	3.596**	3.358***
	(1.33)	(1.56)	(1.23)
Contraception	0.353	-0.391	-0.504
	(0.56)	(0.61)	(0.51)
Smoking during pregnancy	-0.0589	-0.456	-0.351
	(0.65)	(0.75)	(0.62)
Alcohol during pregnancy	0.986*	0.829	0.681
	(0.59)	(0.64)	(0.52)
Marijuana during pregnancy	0.946	-2.120	-1.050
	(1.82)	(1.80)	(1.63)
Cocaine during pregnancy	2.715	7.005***	5.617***
	(2.69)	(2.23)	(2.06)
Constant	93.15***	92.40***	87.06***
	(2.51)	(2.82)	(2.32)
Observations	4623	4623	4623
R^2	0.22	0.17	0.24

Each regression includes the full set of age fixed effects. Coefficients suppressed for readability.
Robust standard errors, clustered at individual child level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: IV Estimates of Motherhood Timing and Scores by Subject

Dependent Variable	(1) Mathematics	(2) Reading recog.	(3) Reading comp.
Mother A1B	0.327**	0.411**	0.189
	(0.16)	(0.16)	(0.15)
Child female	-1.875***	1.076*	0.424
	(0.55)	(0.60)	(0.50)
Child Hispanic	-2.587***	0.594	0.125
	(0.81)	(0.85)	(0.70)
Child Black	-4.224***	-0.597	-2.056***
	(0.85)	(0.89)	(0.72)
Mother AFQT percentile	0.141***	0.140***	0.141***
	(0.016)	(0.016)	(0.014)
Mother HS grad	1.869	2.290	2.361**
	(1.14)	(1.40)	(1.10)
Mother college grad	3.177**	3.155**	3.284***
	(1.35)	(1.58)	(1.25)
Contraception	0.114	-0.889	-0.587
	(0.64)	(0.70)	(0.59)
Smoking during pregnancy	-0.0537	-0.445	-0.349
	(0.65)	(0.75)	(0.62)
Alcohol during pregnancy	0.958	0.770	0.671
	(0.59)	(0.64)	(0.53)
Marijuana during pregnancy	0.894	-2.227	-1.068
	(1.83)	(1.84)	(1.63)
Cocaine during pregnancy	2.674	6.920***	5.602***
	(2.68)	(2.23)	(2.06)
Constant	90.64***	87.20***	86.19***
	(3.91)	(4.13)	(3.68)
Observations	4623	4623	4623
R^2	0.22	0.17	0.24
Hansen's J statistic	2.226	2.036	0.353
P-value	0.329	0.361	0.838

Each regression includes the full set of age fixed effects. Coefficients suppressed for readability.
Robust standard errors, clustered at individual child level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Robustness: Weighted Estimation and Sub-samples for Total Scores

Sample		(1) OLS	(2) IV
All - weighted	Mother A1B	0.444*	0.900*
		(0.23)	(0.49)
	Observations	4623	4623
	R^2	0.22	0.22
	Hansen's J statistic		0.811
	P-value		0.667
Child age 5-8	Mother A1B	0.662***	1.062***
		(0.20)	(0.39)
	Observations	1744	1744
	R^2	0.22	0.22
	Hansen's J statistic		1.876
	P-value		0.392
Child age 5-6	Mother A1B	0.662**	0.982**
		(0.32)	(0.49)
Observations		521	521
R^2		0.26	0.26
	Hansen's J statistic		1.073
	P-value		0.585
Mother A1B $\geq$ 20	Mother A1B	0.630***	0.984**
		(0.22)	(0.42)
	Observations	4474	4474
	R^2	0.26	0.26
	Hansen's J statistic		1.429
	P-value		0.489
Mother A1B $\in$ [20, 29]	Mother A1B	0.680*	1.860*
		(0.37)	(1.01)
	Observations	3525	3525
	R^2	0.26	0.25
	Hansen's J statistic		1.498
	P-value		0.473

Each regression includes the full set of age fixed effects. Coefficients suppressed for readability.
Robust standard errors, clustered at individual child level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Channels for Main Effects on Total Scores: Income

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) IV
Mother A1B	0.432*	0.816*	0.514**	0.868*	0.924**
	(0.26)	(0.46)	(0.22)	(0.46)	(0.41)
Mother A1B* college grad	0.291	0.278			
	(0.39)	(0.72)			
Wage rate in year before birth			-0.0373	-0.0512	
			(0.060)	(0.064)	
Spouse earnings (000s)					0.0795***
					(0.022)
Child female	-0.287	-0.334	-0.552	-0.608	-0.292
	(1.46)	(1.46)	(1.57)	(1.57)	(1.45)
Child Hispanic	-2.094	-1.891	-1.998	-1.834	-1.728
	(2.05)	(2.05)	(2.12)	(2.11)	(2.04)
Child Black	-7.100***	-6.946***	-7.025***	-6.978***	-6.406***
	(2.22)	(2.22)	(2.53)	(2.52)	(2.26)
Mother AFQT percentile	0.433***	0.421***	0.436***	0.425***	0.402***
	(0.039)	(0.040)	(0.040)	(0.042)	(0.040)
Mother HS grad	6.805**	6.628**	0.699	0.449	6.190*
	(3.26)	(3.26)	(4.26)	(4.29)	(3.26)
Mother college grad	2.676	2.338	4.751	4.059	8.668**
	(10.9)	(19.5)	(4.63)	(4.71)	(3.73)
Constant	308.2***	267.8***	310.7***	273.1***	263.4***
	(7.69)	(11.6)	(7.52)	(11.4)	(10.4)
Observations	4623	4623	4019	4019	4623
R^2	0.26	0.26	0.24	0.24	0.26
Hansen's J statistic		1.864		1.118	1.589
P-value		0.761		0.572	0.452

Robust standard errors, clustered at individual child level, in parentheses.
Some coefficients suppressed for readability.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Channels for Main Effects on Total Scores: Family Structure

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Mother A1B	0.639***	1.115**	0.892**	1.040**	1.205**
	(0.23)	(0.44)	(0.41)	(0.48)	(0.52)
Mother married before birth	1.681	0.422			0.702
	(2.83)	(2.90)			(2.94)
Mother never divorced or widowed	3.613**		4.174***		3.156*
	(1.66)		(1.58)		(1.69)
Total children born to mother	0.591			0.928	1.327
	(0.86)			(1.01)	(1.05)
Constant	270.1***	271.0***	272.8***	270.0***	263.5***
	(8.44)	(10.7)	(10.5)	(13.5)	(13.6)
Observations	4192	4192	4623	4623	4192
R^2	0.25	0.25	0.26	0.26	0.25
Hansen's J statistic		1.181	1.025	1.680	1.041
P-value		0.554	0.599	0.432	0.594

Each regression includes the full set of controls. Coefficients suppressed for readability.
Robust standard errors, clustered at individual child level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1