Essays on the Economics of Human Capital

by

Shuqiao Sun

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2020

Doctoral Committee:

Professor Martha J. Bailey, Chair
Assistant Professor Achyuta Adhvaryu
Professor David Lam
Assistant Professor Ana Reynoso

Shuqiao Sun

sqsun@umich.edu

ORCID iD: 0000-0002-2767-7973

For Hongying Qian, Yingwei Sun, and Jingyuan Zhai

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the guidance and help of many others.

First and foremost, I am deeply indebted to Martha Bailey, whose encouragement, patience and wisdom have been invaluable to both my professional and personal growth. Martha inspired my interest in labor economics, demography, and health, and taught me the necessary skills to conduct rigorous work in these areas. Whenever I encountered obstacles in research or in life, I could always count on her unwavering support. She has been the best dissertation committee chair I could hope for.

I am also immensely grateful for having David Lam, Ana Reynoso, and Achyuta Adhvaryu on my dissertation committee. As one of the busiest people on the University of Michigan campus, David was extremely generous to me with his time. Ana has been tremendously helpful in turning me into a confident, well-prepared job market candidate. Ach was a regular source of great academic and professional advice. Together, they gave me the strength I never expected to have found as a doctoral student.

I have been very fortunate to have the opportunity to work with some brilliant co-authors who share my enthusiasm and constantly broaden my perspective. I thank Martha Bailey, Brenden Timpe, Juan Pantano, and Wanchuan Lin for the productive collaboration and insightful discussions we had over the past few years. Writing papers with them is my favorite way to learn economics.

Many faculty members have provided highly constructive feedback on my dissertation. I greatly appreciate the time and knowledge shared by Hoyt Bleakley, John Bound, Charlie Brown, Jing Cai, Sara Heller, Gábor Kézdi, Sarah Miller, Jeff Smith, Michael Mueller-Smith, Frank Stafford, Mel Stephens, and Dean Yang. I owe special thanks to Robert Willis for his mentorship over the years. My fellow graduate students at Michigan provided incredible academic and personal support which I do not take for granted. A non-exhaustive list includes Luis Alejos Marroquin, Ariel Binder, Hannah Bolder, Connor Cole, Thomas Helgerman, Ting Lan, Parag Mahajan, Michael Murto, Kenichi Nagasawa, Dhiren Patki, Elchin Suleymanov, Brenden Timpe, and Wenjian Xu. I also benefited greatly from the experience shared by an older generation of graduate students and postdocs, including Eric Chyn, Valentina Duque, Amelia Hawkins, Alfia Karimova, Morgan Henderson, Daniel Hubbard, Paul Mohnen, and Bryan Stuart. In addition, I thank the organizers and participants of the Michigan H2D2 and labor seminars for making me feel part of an intellectual community.

I would like to express my sincerest gratitude to the Department of Economics for the support I received every semester. The staff members, including but not limited to Mary Braun, Vinnie Veeraraghavan, Laura Flak, Hiba Baghdadi, Lauren Pulay and Julie Heintz, provided excellent administrative assistance. The Institute for Social Research has played an equally important role during my doctoral study. I thank David Lam, David Johnson and Patrick Shields for the attention they have given to my research and study. Especially, I acknowledge the generous support from James N. Morgan and his family, Janet, Salim, Ken and Tim, with the James N. Morgan Innovation in the Analysis of Economic Behavior Fund. As someone who has been using the PSID data since college, I am deeply honored to have received a fellowship named after its founding director. The ISR and PSID staff, including Stephanie Hart, Jennifer Garrett, and Linda Eggenberger, were dedicated to answering all my questions. J. Clint Carter deserves special thanks for the countless hours spent helping me with the restricted Census data.

I have been extremely lucky to have the companionship of many wonderful friends throughout my time in Ann Arbor. I thank Caicai Chen, Xinzhu Chen, Tangren Feng, Xing Guo, Xinwei Ma, Wenting Song, Ruoyan Sun, Yuchen Yang, Xiaoyang Ye, and Hang Yu for always being there for me when I needed them. My life was much brighter thanks to their friendship. During the difficult moments in my journey, their kindness and laughter has been a constant source of comfort and strength.

Finally, I thank my parents, Hongying Qian and Yingwei Sun, and my wife, Jingyuan Zhai. My parents have always taught me to be independent, courageous, and passionate about what I believe in. Even though I am thousands of miles away from home, everything that I am is because of them. Jingyuan is the most thoughtful and caring person I know. I have been blessed with her trust, patience, and unconditional love. Of all the good fortunes I have had, I am most grateful for having her as my partner in life.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

An extensive literature documents the instrumental role of early childhood in improving individuals' life chances. This dissertation consists of three connected chapters that study the effects of early childhood circumstances on human capital formation and its economic consequences.

The first chapter examines the impact of family size on children's long-term wellbeing. The number of siblings is a prominent aspect of childhood family environments that affects parental time and resource investments. Leveraging temporal and county-level variation in access to abortion in the United States during the 1970s, my research design contrasts the adult outcomes of children born just before an abortion clinic became available with the adult outcomes of children born in counties in which abortion remained difficult to obtain. The results suggest that access to abortion decreases the completed number of younger siblings. As their parents avoided unplanned children and achieved smaller family sizes, the children experienced significant improvements in their long-run outcomes, including increased educational attainment, greater labor-force participation, and higher neighborhood quality. The effects also appear to complement the benefits of safety net programs. These findings imply large, persistent returns to reproductive health policies that promote smaller families.

The second chapter, with Martha Bailey and Brenden Timpe, evaluates the long-run effects of Head Start using large-scale, restricted 2000-2013 Census data linked to date and place of birth in the Social Security Administration's Numident file. Using the county-level roll-out of Head Start between 1965 and 1980 and state age-eligibility cutoffs for school entry, we find that childhood participation in Head Start is associated with increases in adult human

capital and economic self-sufficiency, including a 0.29-year increase in schooling, a 2.1-percent increase in high-school completion, an 8.7-percent increase in college enrollment, and a 19-percent increase in college completion. These estimates imply sizable, long-term returns to investing in large-scale preschool programs.

The third chapter, with Wanchuan Lin and Juan Pantano, investigates one underlying mechanism behind the birth order effects on various individual outcomes, with later-born children faring worse than their siblings. We leverage U.S. data on pregnancy intention to study the role of unwanted fertility in the observed birth order patterns. We document that children higher in the birth order are much more likely to be unwanted, in the sense that they were conceived at a time when the family was not planning to have additional children. Being an unwanted child is associated with negative life-cycle outcomes as it implies a disruption in parental plans for optimal human capital investment. We show that the increasing prevalence of unwantedness across birth order explains a substantial part of the documented birth order effects in education and employment. Consistent with this mechanism, we document no birth order effects in families who have more control over their own fertility.

CHAPTER I

Less is More: How Family Size in Childhood Affects Long-Run Human Capital and Economic Opportunity

Abstract

This paper examines the impact of family size on children's long-term wellbeing. The number of siblings is a prominent aspect of childhood family environments that affects parental time and resource investments. Leveraging temporal and county-level variation in access to abortion in the United States during the 1970s, my research design contrasts the adult outcomes of children born just before an abortion clinic became available with the adult outcomes of children born in counties in which abortion remained difficult to obtain. The results suggest that access to abortion decreases the completed number of younger siblings. As their parents avoided unplanned children and achieved smaller family sizes, the children experienced significant improvements in their long-run outcomes, including increased educational attainment, greater labor-force participation, and higher neighborhood quality. The effects also appear to complement the benefits of safety net programs. These findings imply large, persistent returns to reproductive health policies that promote smaller families.

JEL Codes: J13, J24, I2

[T]he shift from couples having large families and making small investments in their children to having small families and making large investments in their children is one of the fundamental dimensions of economic development.

– David Lam, The Population Association of America Presidential Address, 2011

1.1 Introduction

From 1960 to 2017, the average number of children per U.S. woman fell from 3.7 to 1.8.[1] This rapid transition to smaller families is associated strongly with improvements in children's education and economic outcomes. Theoretical work in economics suggests that family size influences parental investment and, therefore, "child quality" (Becker, 1960; Becker and Lewis, 1973; Willis, 1973), as well as economic growth (Becker and Barro, 1988). Consequently, policymakers in many countries have implemented policies to influence parents' decisions regarding the number of children they have. Besides the clearest example from China's One Child Policy, countries such as Mexico, India, Colombia, and Indonesia promote family planning publicly. In the United States, advocates of family planning programs emphasize family size as a crucial determinant of child development.

However, little empirical research has shown that family size has a causal effect on children's lifetime outcomes. The major challenge is the endogeneity of family size, which is potentially correlated with omitted variables.[2] Recent studies rely on instrumental variables, using arguably exogenous sources of variation in family size due to twinning (Rosenzweig and Wolpin, 1980) or the sex composition of existing children (Angrist and Evans, 1998). Yet these pioneering strategies provide remarkably inconclusive evidence on the effects of family

[1] Total fertility rates according to OECD (2019), Fertility rates (indicator). doi: 10.1787/8272fb01-en (Accessed in June 2019)

[2] Numerous studies find that smaller families are associated with better child outcomes, but such descriptive studies are subject to potential omitted variable problems. See Schultz (2005) for a review of these studies.

size on child outcomes (Angrist, Lavy and Schlosser, 2010).[3] A separate strand of literature evaluates the effects of reproductive health policies more generally, but such evaluations do not disentangle the family size effect from the selection effect. For instance, extensive literature documents improved cohort characteristics after abortion legalization, including education and economic outcomes, but attributes such gains largely to the fact that children are more likely to be born into families with more advantageous environments after abortion was legalized. However, what these estimates truly reflect is the combined influence of the selection channel and an effect through family size.[4] More than 50 years after the onset of America's most recent demographic transition, the size *and* sign of the effects of family size on children remain poorly understood.

This paper provides new evidence on the long-run effects of family size by exploiting the staggered roll-out of U.S. abortion clinics. Using information on the number of abortion service providers between 1970 and 1979, I show that abortion service availability varied considerably by county. This staggered introduction at the county level provided an exogenous shock to family size for cohorts born before the roll-out. Using an event-study empirical strategy, I show that children born just before abortion became accessible experienced significant reductions in their number of younger siblings relative to children born where abortion remained difficult to obtain. The effect is stronger when less time has elapsed between the birth and the abortion clinic's launch and not driven by selection on family environments. My results show that the cohort born immediately before abortion became available had a family size that was 0.277 smaller on average compared to cohorts without access to abortion service providers.

[3] This approach is also applied to study the effects of fertility on women. Researchers have documented that smaller families increase labor supply for mothers (Angrist and Evans, 1998; Rosenzweig and Wolpin, 2000).

[4] Extensive evidence documents such selection and characteristics of the affected 'marginal child.' See Gruber, Levine and Staiger (1999) and Ananat et al. (2009). Studies on other policies also experience similar challenges. Bailey, Malkova and McLaren (2018) simulate the role of selection in family planning and use a bounding exercise to estimate the quantitative importance of selection and resource effects.

Using large-scale Census/ACS data that contain a rich set of adult outcomes linked to administrative data that contain date and place of birth, I then examine the long-run wellbeing of children born into smaller families. I find that on average, decreasing family size by one child causes the older child to complete 0.146 more years of schooling, reflecting a 1-percent increase in high school graduation and an 8-percent (2.7 percentage points) increase in college completion. The gains are driven primarily by men, who experienced a 0.242-year gain in schooling and a 13.4-percent (4.4 percentage points) increase in college completion as a result of having one fewer sibling. Gains in human capital also persist as gains in economic self-sufficiency. Men experienced a 1.4-percent increase in labor supply on the extensive margin and a 4.5-percent increase in wages. Further, smaller families led to significantly improved living circumstances in adulthood, measured by a neighborhood quality index that consisted of family income, home ownership, and the share of children in poverty in the census tract of residence. In contrast, abortion access had no effect on children born more than eight years before access, since their chances of having younger siblings were unlikely to be affected.

This paper is first to exploit and validate within-state variation in abortion access as an exogenous source of variation in family size. A long-standing concern in the family size literature is that the available instrumental variables are potentially subject to violations of the exclusion restriction. Twinning is more common in older mothers. Having twins also heightens health risks for the babies *and* mothers and the close spacing between two children is difficult for parents to cope with. The same-sex instrument affects sex-specific cost savings, since existing same-gender children might benefit from *hand-me-down* economies which offset the family size effect (Rosenzweig and Wolpin, 2000). Such omitted factors may have unobserved and unclear effects on later outcomes contribute to the imprecision of existing estimates. My findings suggest that abortion availability has a substantial effect on children's family size. Importantly, abortion availability is not associated with family background variables for these children. Contrary to previous research that suggests lack of a detectable family size effect, I find that the reduction of family size induced by abortion access improves individuals' human capital and economic self-sufficiency over the lifecycle.

My research design is novel in that it focuses on cohorts born just *before* abortion service roll-out, instead of on those born after. A primary conclusion of the existing abortion literature has been that abortion not only changes family size and resources, but also induces selection into parenthood and thus influences the composition of children born after the abortion roll-out (Levine et al., 1996; Gruber, Levine and Staiger, 1999; Malamud, Pop-Eleches and Urquiola, 2016; Bailey, Malkova and McLaren, 2018), which means even if the resources available to a family did not change, selection would lead to better outcomes for the average child. To address this problem, I focus on the cohorts born just before abortion became available. I find that these existing children were not different in terms of the composition of their family backgrounds. The roll-out of abortion providers just after they were born, however, affected their childhood family environment through the number of younger siblings they had. This isolation allows me to provide direct evidence on the sign and magnitude of the family size effect. In addition, focusing on pre-abortion cohorts also allows me to exclude the mechanism of *cohort* size and birth timing because the size of the cohort should remain largely unchanged across regions before abortion roll-out, as was parents' ability to intentionally time pregnancies.

Large-scale administrative data not only provide the precision to detect potentially small effects, but they also contain crucial geo-code information at birth and thus allow within-state analyses. Discussions of existing cross-state studies on abortion legalization have raised a concern of confounding state-level factors during the 1970s, the same time abortion was legalized in the United States. Besides various state characteristics that might be trending differently, the Vietnam War brought significant variation in the hardship deferment for paternity (Bailey and Chyn, 2020). Establishment and subsequent elimination of deferment eligibility for paternity during the late 1960s and early 1970s changed the incentives for married couples to have children. In this paper, any changes at the state level are accounted for by the state-by-cohort fixed effects in my empirical model. The effects of family size on long-run outcomes are identified solely by within-state comparisons.

Besides highlighting the importance of family size in its own right, another contribution of this paper is that I find interactions between family size and public programs. Investigating heterogeneity in the effect, I show that a smaller family size during childhood increased long-run human capital returns to Head Start, a large-scale public preschool program aimed at reducing poverty. A decrease in family size also improved the effect of Food Stamps on economic self-sufficiency.

Family size has received broad attention from researchers and policymakers. This paper is first to assess a causal link running from abortion access to family size, and eventually to individuals' lifecycle outcomes. Overall, my estimates suggest a strong, sustained family size effect on a wide array of long-run outcomes. These findings elucidate the role of reproductive health policies in promoting human capital and reducing poverty.

The rest of the paper proceeds as follows. Section 1.2 introduces a conceptual framework for identification of the family size effect. Section 1.3 summarizes prior research and provides background on access to abortion in the United States. Section 1.4 describes the Census/ACS and administrative data and presents the empirical model. Section 1.5 presents the results. Section 1.6 discusses mechanisms and interaction with safety net programs. Section 1.7 concludes.

1.2 Identifying the Family Size Effect: A Conceptual Framework

Access to abortion may have many influences. In this section, I present a conceptual framework that 1) formalizes the distinction between *family size* and *selection* effects through which abortion legalization might affect cohort characteristics, and 2) motivates an identification strategy aiming to isolate the family size effect. This parameterization illustrates the mechanism and is not imposed during empirical estimation.[5]

1.2.1 Model Overview

The seminal model from Becker (1960), Becker and Lewis (1973), and Willis (1973) highlights parents' endogenous fertility choices and the interrelationship between child quantity and quality. Following the parametrization of this model described by Mogstad and Wiswall (2016), consider a unitary household that chooses its number of children, N, and how much to invest in them to achieve child quality Q. Parents divide their resources between private consumption, C, and investment in their children to maximize their utility, or:

$$U(N, Q, C) = [(\alpha N^{\sigma} + (1-\alpha)Q^{\sigma})^{1/\sigma}]^{\nu} C^{1-\nu}$$

The elasticity of substitution between child quantity and quality is $\frac{1}{1-\sigma}$. The price of child quality, p, which specifies a linear child quality production function, is assumed known to the parents.[6] The budget constraint of a household with income I is represented by $I = C + pQ \cdot N$, and is known to the parents.[7] The quality of all siblings within a family is assumed to be equal, despite the well-documented birth order effects (Black, Devereux and Salvanes, 2005; Lin, Pantano and Sun, 2019), in order to highlight the key mechanism.

Both Q and N are choices, but with any exogenously given N, the locus gives the optimally chosen child quality $Q^{*}(N)$. The family size effect is then thought of as the change in the quality of a child when an exogenous shock increases quantity from N to $N + 1$.

[5] A review of the classic quantity-quality model and its extensions appears in the Appendix.

[6] See Cunha, Elo and Culhane (2013), who relax this assumption.

[7] This model assumes child quality production technology is linear, in which Q represents both the child quality outcome and the parental investment in child quality. A potential extension may assume child outcome Y is a function of parental investment Q.

Consider the possibility of a pregnancy in excess of the desired family size due to a contraception failure. A household with its ideal family size thus has another pregnancy and makes subsequent decisions. For household i with ideal family size N_i^*, it may end up in one of the following three scenarios:

1) No extra pregnancy, in which case the utility is:

$$U(N_i^*, Q(N_i^*))$$

2) An extra pregnancy occurs and results in birth, in which case the utility is:

$$U(N_i^* + 1, Q(N_i^* + 1))$$

3) An extra pregnancy occurs followed by abortion, in which case the unwanted child is avoided and the parents have utility:

$$U(N_i^*, Q(N_i^*)) - A$$

where A represents the cost of abortion, monetary and non-monetary. This cost is known after the additional pregnancy is realized [8].

Note that an implicit assumption of this model is that any contraceptive failure when a household has not yet reached its ideal family size is simply treated as mistimed. It represents a planned pregnancy that will be kept, and does not lead to a permanent, additional birth.[9] It also assumes, for simplicity, that there can only be a maximum of one excess birth. Intuitively, couples become more careful and act to eliminate the possibility of future

[8] Another option is to give the child up for adoption, which, similarly, maintains the existing family size at some cost to the parents.

[9] Suboptimal timing, particularly when pregnancy happens sooner than expected, may also affect parental utility and child outcomes (Nguyen, 2018).

contraceptive failures once they have experienced their first unwanted pregnancy.[10]

Parents choose abortion when:

$$U(N_i^*, Q(N_i^*)) - U(N_i^* + 1, Q(N_i^* + 1)) > A \tag{1.1}$$

that is, when the utility loss due to an extra child is large enough and parents' willingness to avoid it outweighs the cost of abortion.

By definition, 'willingness to avoid' equals the decrease of utility when exceeding the optimal family size by one, which is tied to the parameters in the household utility function. For example, families with greater σ_i in their utility function (i.e., greater elasticity of substitution between quantity and quality) have lower decreases in utility moving from N^* to $N^* + 1$. The intuition is that parents with high elasticity find it easier to compensate for the unexpected change in quantity by adjusting quality, and those with low elasticity are affected more severely and are more likely to choose abortion given the same cost.

Let σ^A represent the cutoff σ such that the willingness to avoid equals A. Anyone with $\sigma < \sigma^A$ decides to have an abortion. A lower A leads to a greater threshold value of σ^A, and therefore more individuals decide to have abortions to terminate their pregnancies.

1.2.2 Model Implication: Family Size Effect and Selection Effect

The roll-out of abortion access can be modeled as an exogenous reduction to A. Legalization of abortion makes it easier to obtain, but there are still several monetary *and* psychic costs associated with terminating one's pregnancy even after abortion is legalized. The roll-out of abortion service providers does not simply eliminate A either, and parents might still find it desirable to give birth to the unwanted child. Nevertheless, a decrease to A effectively

[10] I also assume for simplicity that a household does not have fewer children than it wanted. In reality, this is another possible deviation from fertility expectations that couples experience.

truncates the distribution of σ^A across all households, leading to more abortions and fewer households having unwanted children. One key implication is that abortion legalization affects cohort characteristics through two distinct channels — family size and selection.

Consider a measure of quality for the average child in a cohort. Since the average child represents the mean outcome of the entire cohort, the average child outcome is given by:

$$Y = \int_{\sigma > \sigma^A} Q(N_\sigma) f(\sigma) d\sigma$$

where $f(\sigma)$ is the probability density function of σ.

As abortion becomes increasingly available, the cost of abortion drops from A_1 to A_2. Reduced cost leads to more abortions since a family's willingness to avoid is more likely to exceed the cost, and families have smaller N. Any child from a family that would have found the cost of abortion prohibitively high, but now decides to have an abortion experiences the family size effect:

$$\frac{\Delta Q(N)/\Delta N}{\Delta N/\Delta A}$$

Importantly, this effect is experienced by all children in the family, including the existing children born before optimal family size is reached.

Meanwhile, abortion access also alters which pregnancies become births because families have different reactions to the cost reduction depending on their respective σ. Since some parents decide not to give birth, the average child changes due to the selection effect. As A decreases, σ^A increases and more of the lower end of the σ distribution is truncated. Cohorts born after abortion legalization comprise of fewer 'marginal children' due to selection into childbirth.

Collectively, in the measure of the average child's outcome, $Y = \int_{\sigma > \sigma^A} Q(N_\sigma) f(\sigma) d\sigma$, the change in the cost of an abortion affects two distinct channels — σ^A and N.[11]

An empirical strategy that compares cohorts born before and after abortion legalization measures the total treatment effect, consisting of both the family size and selection effects:

$$T = \int\limits_{\sigma > \sigma^{A_2}} Q(N^{A_2}) f(\sigma) d\sigma - \int\limits_{\sigma > \sigma^{A_1}} Q(N^{A_1}) f(\sigma) d\sigma$$

This total effect can be decomposed into the two channels:

$$\begin{aligned}
T &= \int\limits_{\sigma > \sigma^{A^2}} Q(N^{A_2}) f(\sigma) d\sigma - \int\limits_{\sigma > \sigma^{A^1}} Q(N^{A_1}) f(\sigma) d\sigma \\
&= \int\limits_{\sigma > \sigma^{A^1}} Q(N^{A_2}) f(\sigma) d\sigma - \int\limits_{\sigma > \sigma^{A^1}} Q(N^{A_1}) f(\sigma) d\sigma - \int\limits_{\sigma^{A^1}}^{\sigma^{A^2}} Q(N^{A_2}) f(\sigma) d\sigma \\
&= \underbrace{\int\limits_{\sigma > \sigma^{A^1}} [Q(N^{A_2}) - Q(N^{A_1})] f(\sigma) d\sigma}_{\text{Family size effect}} - \underbrace{\int\limits_{\sigma^{A^1}}^{\sigma^{A^2}} Q(N^{A_2}) f(\sigma) d\sigma}_{\text{Selection effect}}
\end{aligned}$$

1.2.3 Family Size Effect: A Valid Identification Strategy

To identify the family size effect, consider the estimate obtained by comparing children born *before* abortion was legalized between legal and illegal regions. Individuals born before abortion legalization still experience the family size effect since they have different possibilities of having an unwanted younger sibling. The change in quality from $Q(N^{A^1})$ to $Q(N^{A^2})$ persists, whereas the change in composition from σ^{A^1} to σ^{A^2} is zero by construction. The

[11]Many studies empirically document the role of selection. This channel becomes an inevitable identification threat when estimating abortion's family size effect on a given family. Likewise, studies that focus on the selection channel also note the potential biases caused by changing family size. See Gruber, Levine and Staiger (1999).

treatment effect can be written as:

$$T = \int\limits_{\sigma > \sigma^{A1}} Q(N^{A^2}) f(\sigma) d\sigma - \int\limits_{\sigma > \sigma^{A1}} Q(N^{A^1}) f(\sigma) d\sigma$$
$$= \int\limits_{\sigma > \sigma^{A1}} [Q(N^{A^2}) - Q(N^{A^1})] f(\sigma) d\sigma$$

To estimate this effect, it is necessary to examine the Q and N of the average child in a cohort born *before* abortion was legalized in places where their parents had access, in comparison with that of the average child in places in which abortion was illegal or difficult to obtain. The identifying assumption that needs to be satisfied is that legalization, conditional on covariates, must be unrelated to family background and previous child investment decisions. Moreover, legalization can influence the children only through family size. The following sections describe the institutional background of abortion legalization and introduce several validity tests that support this identifying assumption.

1.3 History of Abortion Legalization and Service Roll-Out

1.3.1 Legalization and its Family Influence by State

For nearly three-quarters of the twentieth century, abortion was illegal everywhere in the United States. In 1970, five states became the earliest to effectively legalize abortion, when New York, Washington, Alaska, and Hawaii repealed anti-abortion laws, and California entered an era of de-facto legalization following a late-1969 ruling that abortion ban was unconstitutional (Potts, Diggory and Peel, 1977). In January 1973, abortion became broadly available in all states, following the landmark decision in *Roe v. Wade*.

The impact of legalization was both immediate and large; the number of abortions increased sharply. The monetary cost of the procedure fell significantly to $80 from the previous $400 to $500 for illegal services (Kaplan, 1988) or the cost of traveling abroad (often to Europe) while pregnant. As a result, the U.S. birth rate dropped considerably during the early 1970s (Levine et al., 1996; Levine, 2004; Angrist and Evans, 2000). Levine et al. (1996) estimate an 8 percent reduction in the birth rate following the actions of the early-legalizing states. Their primary difference-in-difference results suggest a decline in relative birth rates between early-legalizing states and the remainder of the country starting in 1970, which reversed when all states legalized abortion beginning 1973. Notably, the birth rate decline discussed here is contemporaneous. It remains unclear whether this dramatic change in birth rate represents a decrease in women's completed fertility.[12]

Using a longitudinal survey dataset from the Panel Study of Income Dynamics (PSID), I first examine the impact of abortion legalization on women's completed fertility. The dataset was constructed at the child level and contains individuals' years of birth, states of birth, and numbers of siblings during their lifetime, obtained from linked mothers' birth records in the PSID Childbirth and Adoption History file.[13] These variables allowed me to examine whether abortion legalization led to smaller family sizes for the existing children.

Both demography and economics literature evinces the effectiveness of abortion on a family's ability to avoid unwanted childbirths. For existing older children who live in such families, it follows that they experience decreases in the number of younger siblings. Indirect evidence suggests that this effect is strong; most unwanted children in the United States are higher order births, who would have been younger siblings to children in the same family (Child Trends, 2013; Lin, Pantano and Sun, 2019).[14]

[12] One assessment of this topic can be found in Ananat, Gruber and Levine (2007), who document an increased number of childless women. Due to limited data, they document no completed fertility effect at the intensive margin, or influences on affected cohorts' long-run wellbeing.

[13] To ensure family size information reflected the *eventual* number of siblings, I included only children whose mothers most recently reported her number of children ever had after she turned 40. See Appendix A.1 for detailed descriptions of the PSID data and sample construction.

[14] Unwanted children are rare among first births. Parents often consider early-born children mistimed and unintended but nevertheless wanted. Related terminology is well-developed in demography literature. See

Panel A of Figure 1.1 plots the covariate-adjusted difference in family size between early-legalizing states and the remainder of United States by birth year.[15] The pattern is largely as expected, similar to what Levine et al. (1996) document for the birth rate. A reduction in average family size is visually evident after 1967 in early-legalizing states. A gap appears and then closes within three birth cohorts. Important to this paper's identification strategy, the decline first appears among cohorts born two years before (in 1968) the abortion ban was lifted in early-legalizing states (in 1970). Cohorts born further before the repeal had more time elapsed between birth and access to abortion, and therefore had as many younger siblings as when abortion was illegal. The difference remains statistically indistinguishable from zero in every year before 1968, whereas cohort born immediately before legalization experienced the largest reduction in family size (0.232).

Family size depends on many factors, with one concern being that trends related to a mother's behaviors or characteristics might have changed around the same time abortion access changed. Appendix Figure A.1 shows the difference in the number of siblings between early-legalizing and remaining states, stratified by several maternal characteristics. These characteristics are common predictors of unwanted pregnancies and thus relate closely to the probability of having an abortion. Indeed, all of these predictors appear to drive the reduction in the number of siblings more than the overall sample; the effect appears strongest among children born to unmarried women (significant at the 10% level), women over 35, and women with lower incomes or who were non-white.

Panels B and C of Figure 1.1 show regression coefficients by parity. Patterns among first

---

Santelli et al. (2003).

[15] The covariate-adjustment exercise was conducted using generalized difference-in-differences specification: $Y_{i,b,s} = \delta_s + x_{i,b,s}\beta + \phi_t \sum_{j=a}^{A} 1(b=j) + \pi_t \sum_{j=a}^{A} 1(b=j) \times Repeal_s + \epsilon_{i,b,s}$, where $Y_{i,b,s}$ is the total number of younger siblings of individual i born in state s in year b. Vector $x_{i,b,s}$ is a set of individual- and family-level covariates, including race, gender, and mother's completed education and age at birth. δ_s is either a set of state of birth fixed effects or an indicator of whether the state of birth was among early legalizers that repealed the abortion ban in 1970. The coefficient of interest, π_t reflects the change in the relationship between birth cohorts and their eventual number of younger siblings when early-repeal states legalized abortion and other states maintained the ban.

births and higher parities appear similar. The consistency indicates the events that might have created incentives for families to have their first child earlier either did not have a significant impact or the effect is uncorrelated with having a smaller family. Appendix Table A.1 shows the magnitude of the effect by parity and with estimates placed into three-cohort bins to increase precision.

1.3.2 Discussion of Cross-State Analysis

Since the innovative use of the staggered legalization of abortion for quasi-experimental research, many studies have examined the impacts on children and also discussed problems with relying on cross-state variation extensively. Cohorts born after abortion access experience lower infant mortality (Gruber, Levine and Staiger, 1999) and decreased instances of adolescent substance use (Charles and Stephens, 2006). Women who would have mothered these children experience significant increases in college graduation (Ananat et al., 2009). Yet these compelling results are debated constantly (Donohue and Levitt, 2004; Dills and Miron, 2006; Foote and Goetz, 2008; Donohue and Levitt, 2008).

One major challenge is that cross-state comparisons of adolescent and adult outcomes are potentially subject to influences from factors other than abortion. States that repealed abortion bans prior to *Roe v. Wade* are special in that their social policies are progressive in other ways. Laws passed around the same time by these states might correlate with the legalization of abortion. Some of these laws are specifically related to reproductive health. For example, legal restrictions on the birth control pill varied considerably during the 1960s (Bailey, 2010) and were decided largely by states.[16] Other laws did not target family planning but created alternative channels that might affect later-life outcomes considerably. For example, state supreme courts ordered school finance reforms beginning in 1971 (Hoxby, 2001), overturning public K-12 school finance systems for the same cohorts affected most

[16]Bailey (2010) find that sales bans of the Pill does not correlate with repeal of abortion bans.

by abortion legalization. Claims that abortion legalization lowered crime rate are disputed by Joyce (2004) who notes the potentially confounding crack cocaine epidemic, which began during the late 1980s and arrived in New York and California earlier than elsewhere.

Even when changes occurred at the national level, they threaten identification in situations where they cause heterogeneous effects across states. Bailey and Chyn (2020) found a considerable fertility response during the Vietnam War when American families sought ways to avoid military service; one way to qualify for deferment was having children. The incentive appears significantly larger in the five states that legalized abortion earlier, since their anti-war sentiment was much higher than elsewhere in the United States. This differential response creates confounding factors for abortion estimates that resulted from cross-state comparisons, since children born to families that had compelling reasons to avoid serving in the military might have grown up in very different childhood circumstances than others did.

To circumvent these concerns about cross-state analysis, this paper exploits the staggered availability of abortion clinics at the county level.[17]

1.3.3 County-level Roll-Out of Abortion Services

Although abortion was legalized by states and then nationwide later, its availability varied considerably across the country. As many critiques of state-level comparisons point out, an

[17] As a proof of concept, I also implement a similar design as that used in previous cross-state comparisons. Appendix Table A.2 presents state-level estimates on children's completed education using PSID. The results suggest that those born immediately before the 1970 policy change in early-repeal states tended to complete more years of education in comparison to older cohorts, and in comparison to individuals born in non-repeal states. In the preferred specification with year and state of birth fixed effects (column 2), consistent with the hypothesis that having unwanted younger siblings causes disruptions in parental investment and compromises family environments, the same cohorts of individuals who experienced a decrease in young siblings also experienced a 0.321-year increase in completed education. The effects appear driven by higher education, with college completion having had the largest effect (15.5 percentage points, column 8). If there were state-level changes, state-specific trends, or national events that imposed differential effects that correlated positively with both repeal of the abortion ban and education, then these estimates represent the upper bounds of abortion's human capital effects through family size.

ideal research design would be to exploit some form of within-state variation of abortion access (Foote and Goetz, 2008).

Using county-level information on service providers made available by the Guttmacher Institute,[18] Figure 1.2 shows the roll-out of abortion service providers in each continental U.S. county.[19] The map shows idiosyncratic variation in access timing; some counties were able to establish service access immediately, likely converting existing clinics to part-time providers. Others did not have service providers until much later. The staggered introduction is widely observed within each state.

Demonstrating the validity of a roll-out design is critical to this paper's analysis. A formal test would need to show that the year abortion services began in a county is unrelated to various county-level characteristics, and to potential underlying trends in fertility that are county-specific. One hypothesis is that abortion clinics first appeared in areas that are more progressive, highly-educated, and affluent. Economic circumstances are commonly associated with fertility declines (Foster, Rosenzweig et al., 2007), and areas might have varying demand for abortion based on education and income level. This would have likely caused some areas to gain access to abortion earlier and be more developed in human capital at the same time. Policymakers and service providers, and their funding sources, might also have targeted some areas. For example, resources might be directed first to areas with the highest incidence of unwanted births or observe an increasing trend. All of these channels create potential violations of the exclusion restriction, which requires abortion service roll-out to be uncorrelated with unobserved determinants of child development.

I examine such potential correlations in Table 1.1. I regress a set of county characteristics, collected from the 1970 County and City Data book, on the timing of first abortion service providers in counties that obtained access between 1970 and 1979. Among 16 characteristics,

[18]This information was cleaned and used by Bailey (2012)

[19]Information on the number of service providers in New York, California, and Washington between 1970 and 1972 is missing and was therefore extrapolated from 1973.

the only predictors of getting a service provider early during the 1970s are total population and share of urban population. It appears, as one might expect, that the service arrived first in the most populous places and thus those most in need. The demographic pattern motivates use of urban-population-by-cohort time trends, which I control for in all empirical models.[20]

Notably, roll-out of abortion service providers does not appear to correlate with income, education, leading political party, or the age structure of a county's population. Service seems to have begun in both affluent and less-developed counties following the legalization, despite concerns about endogeneity arising from abortion's varying popularity. A more detailed comparison of different groups of counties appears in the Appendix.

Another threat to identification is that timing of abortion clinic's roll-out might be correlated with counties' fertility rates, which causes omitted variable bias because fertility rate is associated with an area's many unobserved characteristics that also determine children's development. Figure 1.3 examines this correlation directly. Using data from the NCHS Natality Detail Files, panel A plots county-level general fertility rates on the timing of abortion clinic's roll-out. It excludes counties from early-legalizing states and focuses on the period between 1973 and 1979. Roll-out timing does not appear to be correlated with the *level* of fertility rates immediately before *Roe v. Wade*.

Exploiting the staggered timing also requires careful examination of pre-trends. If clinics were funded to specifically target areas in which fertility rates remained high despite the national trend, were trending up, or were decreasing at a slower rate than the national average, then this identification strategy might falsely attribute these trends to abortion availability. Panel B, Figure 1.3 plots changes in fertility rates between 1968 and 1972 and the timing of abortion clinic's roll-out. A proactively targeted roll-out scheme would predict a negatively

[20] Share of urban population, u, was used to generate five categorical variables, indicating $u = 0$, $0 < u \leq 25$, $25 < u \leq 50$, $50 < u \leq 75$, $75 < u \leq 100$, and interacted with cohort trends.

slopped line fitted through the figure, but the actual pattern suggests no evidence of selective roll-out at the county level. Despite various hypotheses regarding how service might appear, actual roll-out of abortion service providers appears largely idiosyncratic.

1.4 Census, Administrative Data, and Research Design

1.4.1 The Data

An ideal dataset for analyzing the long-run impact of childhood family size requires information on 1) place of birth or place of residence in early childhood, 2) human capital and economic outcomes well into adulthood, and 3) a sample large enough to detect potentially small effects and that contains enough counties. This paper relies on large-scale, restricted-access data to achieve identification.

I use the newly available long-form 2000 Decennial Census and the 2001-2016 American Community Surveys (ACS) linked to the Social Security Administration's Numident file through a protected identification key (PIK). The Census/ACS data include nearly one-quarter of the U.S. population and observe an extensive set of adult outcomes on educational attainment, labor-market participation, family income and poverty status, as well as living circumstances and neighborhood quality.

To focus around the roll-out of abortion services, my sample is comprised of individuals born between 1960 and 1986 and individuals between age 25 and 54 in their prime-earning years. The sample excludes individuals with allocated or missing values and counties with unknown numbers of abortion clinics. To minimize disclosure burden from the Federal Statistical Research Data Center, I constructed a full-information sample to further exclude individuals who are missing any of the outcome variables, except for outcomes that are logged. The result is a final analysis sample of about fifteen million American adults.

Linkage to the Numident file provides information on individuals' exact places of birth, which are matched to counties of birth (Stuart, Taylor and Bailey, 2016). The dataset is then merged with county-level abortion service provider information to measure precisely the relative time between an individual's year of birth and the year of local service access. One shortcoming is lack of family background information in the data, which might help identify mechanisms or create a high-impact sample. Nevertheless, findings using PSID suggested a strong family size channel that can be plausibly isolated, and the sample size is sufficiently large to detect important effects that might be small in magnitude.

To minimize computational resources, I collapsed the data into cells by birth year, survey year, and county of birth. When used during analyses, the cells were weighted using the number of observations in each cell (Solon, Haider and Wooldridge, 2015). I also constructed indices by normalizing and grouping long-run outcomes into three categories — human capital, economic self-sufficiency, and living quality — which helped increase statistical power and mitigated issues related to multiple hypothesis testing (Kling, Liebman and Katz, 2007).

1.4.2 Research Design to Capture Abortion's Impact

My research design exploits the natural experiment of abortion service providers' roll-out from 1970 to 1979, with a flexible event-study framework,

$$Y_{bct} = \theta_c + \alpha_t + \delta_{s(c)b} + \boldsymbol{Z'_{cb}}\boldsymbol{\beta} + \sum_y \pi_y 1\{b - T_c = y\} Abortion_c + \epsilon_{bct} \tag{1.2}$$

where Y_{bct} is a measure of adult human capital, self-sufficiency, or living quality at time of survey t for individual born in county c in year b. θ_c is a set of county fixed effects, α_t is a set of survey year fixed effects, and $\delta_{s(c)b}$ is a set of state-by-cohort fixed effects. Importantly, controlling for state-by-cohort effects alleviates concerns raised by critics regarding changes to state policies and varying prior trends that affected birth cohort differentially. Also notably, this level of flexibility accounts for changes in travel distance to one of the early-legalizing states, a factor highlighted by Joyce, Tan and Zhang (2013) and that is also captured by state-by-cohort effects.

To control for several time-varying characteristics at the county level, I include $\boldsymbol{Z'_{cb}\beta}$, a set of county-level observables and demographic controls measured in an individual's county and year of birth (see also Hoynes, Page and Stevens 2011; Bailey 2012; Bailey and Goodman-Bacon 2015). These observables are interacted with a linear trend in year of birth, except for a set of categorical indicators for county population, which is interacted with year of birth dummies to allow more flexible control of the changing population density. $Abortion_c$ is an indicator for ever having at least one abortion service provider between 1970 and 1979. All standard errors are corrected for heteroskedasticity and adjusted for an arbitrary within-county covariance structure (Arellano, 1987; Bertrand, Duflo and Mullainathan, 2004).

T_c denotes the year abortion first became available in county c. Event-time y represents an individual's birth year relative to the local roll-out of abortion services. The point estimates of interest are π_y, which capture the evolution, by event-time, of the educational or economic outcomes in counties with access net of changes in untreated counties. The omitted category is set to -8. For example, if the first abortion clinic increased human capital significantly for the cohort born two years before access (i.e., $t = -2$) in comparison to those born eight years before access, the point estimate for π_{-2} should be positive and statistically significant.

The expected results are that for the cohort born immediately before access became available ($t = -1$), the effect on educational and economic outcomes through decreasing numbers of younger sibling should be strongest. Individuals born after abortion became available ($t >= 0$) experienced a combination of the family size and selection effects. Abortion should have little effect on cohorts born earlier ($t <= -8$). Effects from π_{-8} to π_{-1} should increase gradually. Based on the estimated first-stage effect on family size, a test of joint significance for $\pi_{-3,-2,-1}$ should be positive and statistically distinguishable from zero.

1.5 How Abortion Impacts Family Size and Long-Run Wellbeing

1.5.1 Abortion Access and Family Size by County

I first examine the effect of abortion providers' roll-out on family size. Information on individuals' date and county of birth is obtained from the restricted-use geo-coded PSID, which I linked to abortion service providers data. I also linked the main PSID interviews with the PSID Childbirth and Adoption History file to obtain information on siblings. Appendix A.2 describes the PSID data and sample construction.

I tested the primary empirical specification with several modifications. First, the dependent variable is the completed number of siblings for individual i born in county c in year b. To ensure the sibling count describes an individual's *completed* family size, I collected information only when the most recent observation of the mother was after she turned 40. Second, to improve the precision of the estimates, I controlled for a set of individual-level covariates, including gender, race, maternal education, and birth order in family. I also grouped event-time dummies into three-year bins.

Figure 1.4 presents the event-study estimates of this first-stage effect. Compared to a control group of children born approximately eight years before an abortion service provider first appeared in their county, children born closer to abortion becoming available experienced significant reductions in their completed number of siblings, averaging 0.277 fewer siblings (statistically significant at the 10-percent level) for children born the year before the service launched. In contrast, the change in completed family size remained indistinguishable from zero for individuals born eight years or more before abortion became available, since the roll-out did not arrive soon enough to affect younger siblings born after them.

An alternative identification strategy to the event-study specification above is to exploit cross-county variation in the intensity of providers. I replaced the set of event-time dummies, $\sum_y \pi_y 1\{b - T_c = y\}Abortion_c$ in equation (2), with a continuous measure of the number of abortion service providers per 1,000 women aged 15 to 44. This intensity measure is denoted $1(b = j)Prov_c^{73-78}$, which describes the average number of abortion service providers in a county within 5 years of *Roe v. Wade* (1973 to 1978) per 1,000 women aged 15 to 44 interacted with indicators of birth cohorts. Since information on abortion service providers is missing from 1970 to 1972, this exercise also excluded the five early-legalizing states from the sample.

Table 1.2 presents estimates from this service intensity analysis. For cohorts born just before 1973 in counties with more access to abortion immediately after *Roe v. Wade*, their completed family size experienced a significant decrease. The specification that includes state-specific trends and county characteristics interacted with linear time trends (Column 2) accounts for differential trends that might confound cross-sectional analysis. The 1970-1972 birth cohorts experienced a decrease in the number of younger siblings by 3.18 per increase in abortion service providers per 1,000 women aged 15 to 44, in comparison to the omitted category of 1960-1963 cohorts. Consistent with the event-study estimates in Figure 1.2, children born more than three years before *Roe v. Wade* do not display any negative relationships between family size and access to abortion. The magnitude of the estimate for 1970 to 1972 cohorts, when multiplied by the sample's average abortion service provider intensity of 0.07, implies an estimated decrease of 0.223 siblings. These results provide further evidence that within-state variation in the availability of abortion is a determinant of children's family size.

1.5.2 Long-Run Effects on Human Capital, Self-Sufficiency, and Living Quality

I next implement the event-study specification to examine the effect of fewer younger siblings on adult outcomes of these children using three indices — human capital, economic self-sufficiency, and neighborhood quality.

Figure 1.5 plots estimates for the overall sample in event-time. Among compositionally balanced birth cohorts (i.e., individuals born between fourteen years before and five years after the roll-out of the first abortion clinic), the patterns appear as expected. For each index, access to abortion service had no measurable effect on children born eight years or more before. For these children, it is likely that their parents gave birth to them and their younger siblings while abortion was still illegal or difficult to obtain. However, estimates for all three indices exhibit an increase between event time -8 and -1, when children were born closer to the year abortion service became available. The magnitudes of the effects appear particularly large when approaching event-time -1 (i.e., born just before abortion roll-out), and the joint test of $\pi_{-3,-2,-1}$ rejects the hypothesis of no effect on the human capital and neighborhood quality indices. To the right of event-time -1, the patterns become less consistent across outcomes. Indeed, in the case of economic self-sufficiency and neighborhood quality, estimates become larger than those at -1, potentially due to selection into childbirth.

Event-study estimates plotted in Figure 1.5 are intent-to-treat (ITT) effects. Table 1.3 presents the average treatment-effects-on-the-treated (ATET) by scaling event-time magnitudes using this paper's estimated effect on completed family size, a reduction of 0.277. To investigate which characteristics drive the increase in the indices, Table 1.3 also summarizes the estimates for the main components that comprise three indices. Column 1 presents the mean of the outcome for the control group; individuals born 8 years before abortion roll-out.

Column 2 summarizes ITT estimates and standard errors for each outcome of children born just before abortion roll-out (event-time -1). Column 5 presents ATET (coefficients in Column 2 divided by 0.277). Column 6 presents the percentage change implied by the ATET relative to the control group (the ratio of Column 5 to Column 1). I also include results from formal joint tests of zero effect on children born just before roll-out (i.e., event-time -3 to -1, Column 3), and on children too early to be affected by abortion access (i.e., event-time smaller than -8, Column 4). The total number of observations $(15,891,000)$ highlights the large sample size due to the restricted Census/ACS and administrative Numident data.

Overall, a decrease in family size by one caused the standardized human capital index to increase by 4.2 percent of a standard deviation, reflecting a 0.146-year increase in completed education. A major advantage of the large-scale data, besides improved precision, is the ability to investigate a rich set of outcomes that cover many dimensions well into children's adulthood. Estimates for the individual components of the human capital index suggest a moderate 1-percent increase in high school completion and a much larger increase in college completion of 8 percent (2.7 percentage points). College education is the costliest child investment in the United States, and the large estimate accords with the hypothesis that family size mostly affects whether parents can afford to send their children to college. The estimate also appears large for having a professional or doctoral degree, another expensive investment that might depend heavily on families' credit constraints. Individuals born 8 years or more before abortion became available experienced no effect since their family sizes were unlikely to have decreased. With the exception of professional or doctoral degree, tests of joint significance for $\pi_{y,y<-8}$ fail to reject the null hypothesis of no effect in all outcomes.

Since family environments might affect boys and girls differently, Table 1.4 presents estimates by gender. Gains in human capital appear driven by men, who experienced a 0.242-year gain in schooling, a 1.6-percent (1.5 percentage points) increase in high school completion, and a 13.4-percent (4.4 percentage points) increase in college completion. Having small families also significantly increased men's likelihood to obtain a professional or doctoral degree (1 percentage point) and work in a professional occupation (3.1 percentage points). Overall, for men who had one fewer sibling, the human capital index increased by 6.8 percent of a standard deviation (statistically significant at the 1-percent level).

Gains in male human capital also persisted as gains in economic self-sufficiency. On average, men experienced a 1.4-percent increase in labor supply (Table 1.4, column 6) on the extensive margin and a 1.5-percent increase in weeks of work. They also experienced a 4.5-percent gain in wage income, making them 17.6-percent less likely to live in poverty. In contrast, the effect of family size on girls' long-run self-sufficiency is imprecise despite gains experienced in years of schooling (0.05 year) and professional or doctoral degree completion (32.3 percent, 0.9 percentage points). Women's labor-force participation after growing up in small families appears to have *decreased*. This reduction might be surprising, but is consistent with an income effect resulted from marrying higher-earning men. This channel remains to be tested since one important dimension missing from this exercise is direct observation of family structure and marriage quality.

Nevertheless, estimates for a group of census tract characteristics provide additional evidence of how men and women experienced changes to their adult living standards. Panel C, Table 1.3 and Table 1.5 presents results on neighborhood quality measures. For females and males, children born just before abortion access experienced significant improvements in measures of quality of their census tract of residence in adulthood. On average, their neighborhoods have nearly 2-percent more home ownership (significant for women and marginally significant overall).[21] Smaller family also caused a 4.5-percent increase in adult neighborhood income and a 6.9-percent decrease in share of children in poverty in one's neighborhood. The summary index of neighborhood quality experienced an increase of 4.8 standard deviation when

[21]Corrections for multiple hypothesis testing is needed to interpret the p-values for individual outcomes. An exercise to account for multiple hypothesis testing appears in the Appendix.

decreasing family size by one.

1.5.3 Testing for Confounding Changes

The main empirical strategy implemented in this paper is valid only if the roll-out of abortion service providers is uncorrelated with other factors that might affect children's long-run outcomes besides family size. In addition to validity checks in Section 1.3, this section investigates a variety of family background variables to test for changes in determinants of childhood environment that could arise in the treatment group in comparison to the control group. Specifically, my research design, which focuses on the cohorts born before abortion service providers became available, aims to circumvent the issue of selection causing compositional changes of children. Ruling out changes in average family-level observables provides compelling evidence that results are not driven by selection into childbirth.

To examine these confounding factors, Figure 1.6 applies the specification in equation (2), but replaces the dependent variable with a series of family-level observables. To examine whether sample selection happens to a certain parity, it also stratifies by first and higher-order births. The results show no detectable effects on maternal education, mother's marital status or age at birth, mother's age at birth, or probability of low birth weight.[22] The effects on number of siblings is significant despite the non-results for these observable characteristics, which supports my hypothesis that for the cohorts born immediately before abortion legalization, the policy change affect their human capital *only* through an improvement in childhood family environment facilitated by a better planned family.

The concern this research design aims to address is that identification using cohorts born after abortion legalization will be confounded by changing composition of the sample. Figure 1.6 panel B shows this concern appears to be real. For the cohorts born after abortion

[22] Parents may establish fertility stopping rules. For instance, they might decide against having more children if a child has low birth weight.

became available, their parents appear to be more educated, older, and more likely to be married. It is also suggestive that they are less likely to have low birth weight. Therefore, even if significant improvements in life cycle outcomes can be observed for these cohorts, there seems to be little reason to attribute these impact solely to improved childhood family environment.

1.6 Interactive Effects: Complementarity or Substitutability?

A long-standing hypothesis about human capital intervention programs is that investing in early childhood generates substantial long-term returns by increasing the returns to subsequent investments. Disadvantages in early childhood grow over time (Currie and Thomas, 2000), and interventions at this stage are effective in breaking the cycle of poverty because they increase the returns to interventions at later stages; a feature formalized by Cunha and Heckman (2007) as dynamic complementarity. Although it is impossible to pin down the exact stage when family size intervenes during the dynamic process of child development, evidence of complementarity or substitutability with other human capital programs provides useful insights for evaluating such programs as well as understanding the mechanisms of family size effect. This section examines interactive effects with several public programs introduced in the 1960s. These programs aimed to provide social safety net for disadvantaged children and tackled the causes poverty from several aspects. Exposure to these programs varied considerably in late 1960s and early 1970s across counties (Bailey and Duquette, 2014).

I use the Census/ACS data and implement an extension of the county-level event-study specification. In addition to the main effects of event-time dummies, I include terms that capture the interaction between each event-time dummy and an indicator of high exposure to each public program (Head Start, Food Stamps, Medicaid, and CAP[23]), $\sum_y \phi_y 1\{b - T_c = y\} Abortion_c \times HiExp_{bc}$, to test hypotheses of effect interactions.

The first hypothesis is that reduction in family size facilitates more parental time which complement other programs. Children from smaller families receive more help after school, more attention when having health conditions, and in general are better prepared to benefit of human capital interventions during their childhood. Table 1.6 suggests that smaller family size complements both Head Start and Food Stamp Programs. The effects of preschool provided by Head Start on human capital and neighborhood quality are significantly larger in places where children experienced reduced family sizes. The Food Stamp Program has a significantly positive interaction with smaller family size on economic self-sufficiency outcomes. It appears that children in smaller families benefit more from the support of Food Stamps. These findings suggest evidence of complementarities between reduction in family sizes and two of America's largest safety net programs.

Another hypothesis it that reduction in family size and some programs are substitutes. One would expect a negative interaction if the decrease in family size help families avoid experiencing hardship that are targeted by some programs. For example, Gillezeau (2010) documents that the Community Action Program (CAP) decreased the number of riots in 1960s and 1970s. These programs might not have significant impacts if families are already living in peaceful neighborhoods. Consistent with this hypothesis, the CAP's effect on neighborhood quality, which includes average family income and property value in individual's census tract, is significantly smaller for children from smaller families. There also appears to be a negative albeit statistically insignificant interaction between the effects of CAP and smaller families on human capital and self-sufficiency.

Additionally, Appendix Table A.4 uses the PSID data and shows the evidence also exists on

[23] Information about Head Start grants is collected from the National Archives Community Action Program electronic files and made available through the work of Bailey and Goodman-Bacon (2015) and Bailey, Sun and Timpe (2018).

the state level.[24] Among individuals born between 1960 and 1980, Head Start has a positive and significant effect on completed years of education. Moreover, having an abortion-induced reduction in family size, indicated by born before 1970 in an early-repeal state, increases Head Start's human capital benefit by 1.02 years of education. The effect of the interaction term is largely driven by college attendance, suggesting that the monetary cost of attending college could be a key factor when families with Head Start access decide their children's education based on family size. On the other hand, college completion rate responds only to the family size treatment (and is marginally significant for the Head Start treatment), but not to the interaction between the two treatments.

1.7 Concluding Remarks

Since Becker (1960), Becker and Lewis (1973) and Willis (1973) formalized the relationship between child quantity and quality in theory, many researchers have attempted to identify the family size effect in practice, and their findings have been mixed. Especially with modern identification strategies and more comprehensive data from developed countries, estimates have been remarkably imprecise and inconsistent with the theoretical predictions. The endogenous nature of choosing one's family size creates potential threats to identification, and most instrumental variables used previously might affect unobserved factors in the complicated process of raising children. Many researchers attribute the lack of evidence to these empirical caveats.

This paper provides new evidence on the impact of family size. The source of variation I exploit is the staggered county-level roll-out of access to abortion service providers. For

[24] In addition to the terms specified earlier, this model includes an interaction between accesses to abortion immediately after birth, and access to Head Start in county of birth when the individual is age eligible (age 5 or younger), $1(b = j) \times Repeal_s \times HeadStart_{bc}$. Additionally, x_{bct} includes interactions between Head Start access and cohorts, and the Head Start indicator. The result is a fully interacted model between cohorts, an indicator of early-repeal state, and an indicator of being age-eligible when Head Start first begun in the individual's county of birth.

the first time, I document that significant reductions in completed family size occurred for children born just before abortion became available; their parents became more capable of avoiding unplanned children and thus had smaller families in comparison to those who still could not obtain abortion services easily. To examine long-run outcomes, I use confidential Census/ACS data linked to the county and date of birth information in the large-scale administrative Numident file. I find that reductions in family size generated substantial long-run improvements in human capital, economic self-sufficiency, and neighborhood quality. Decreasing family size by one led to an 8-percent (2.7 percentage points) increase in college completion, a 1.4-percent (1.3 percentage points) increase in men's labor supply, and a 4.5 log-point increase in adult wages. For both men and women, smaller families led to better living quality in adulthood, represented by significant increases in average home ownership and average family income of the neighborhood in which they live.

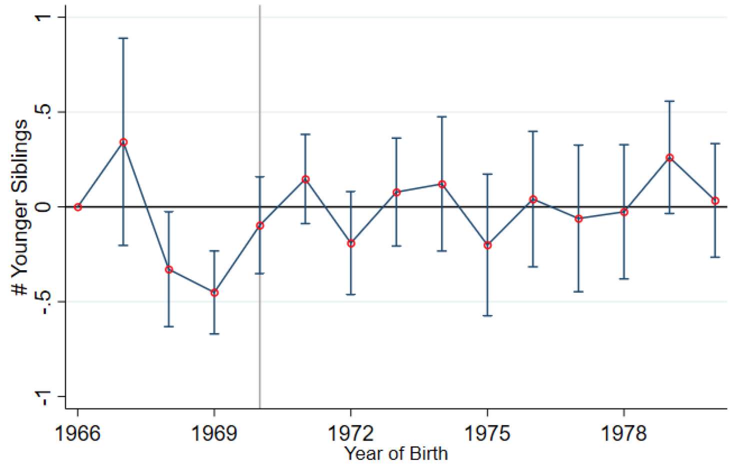
My estimates came from cohorts born just before abortion roll-out instead of those born after. Literature on abortion legalization documents that abortion leads to selection. Children born after abortion access were selected into birth and came from different family backgrounds on average. Focusing on pre-abortion cohorts circumvents this problem and isolates the causal effects of decreased family size. I find no change in measures of family background among pre-abortion children; abortion access affected them only through family size. To this front, my findings also contribute to literature that evaluates consequences of abortion legalization, which commonly highlights a selection effect on the marginal child. My findings suggest that access to abortion also had substantial impacts on the 'inframarginal' children; the children already born.

It is worth noting that the existing quantity-quality literature provides an extensive discussion on why evidence on the family size effect remains inconclusive. Apart from the concern regarding the validity of existing instruments, families with more children might genuinely benefit from economies of scale or have a positive interaction between child quantity and quality in their preferences. Further, any particular instrumental variable can identify only effects on a group of individuals affected by the instrument (Imbens and Angrist, 1994; Angrist, Lavy and Schlosser, 2010). Effects might vary considerably on different margins, and in some cases might even change sign. This paper not only introduces a plausibly exogenous variation in family size that induces substantial effects, but focuses on the margin that aligns with the target of many reproductive health policies. I argue that a smaller family matters, particularly when it is achieved by avoiding having unplanned children in excess of a family's desired fertility.

In regards to implication for human capital and anti-poverty programs, this paper also demonstrates interactions between family size and social safety net programs aimed to support disadvantaged children. My findings imply significant policy complementarities; having small families increases long-run human capital returns to Head Start and the economic impacts of Food Stamps. Overall, these results elucidate how reproductive health policies can, by reducing family size, facilitate children's human capital formation, increase their economic opportunities, and help families escape poverty.

Figure 1.1. Difference in Number of Siblings, Early-Legalizing vs. Other States

A. All

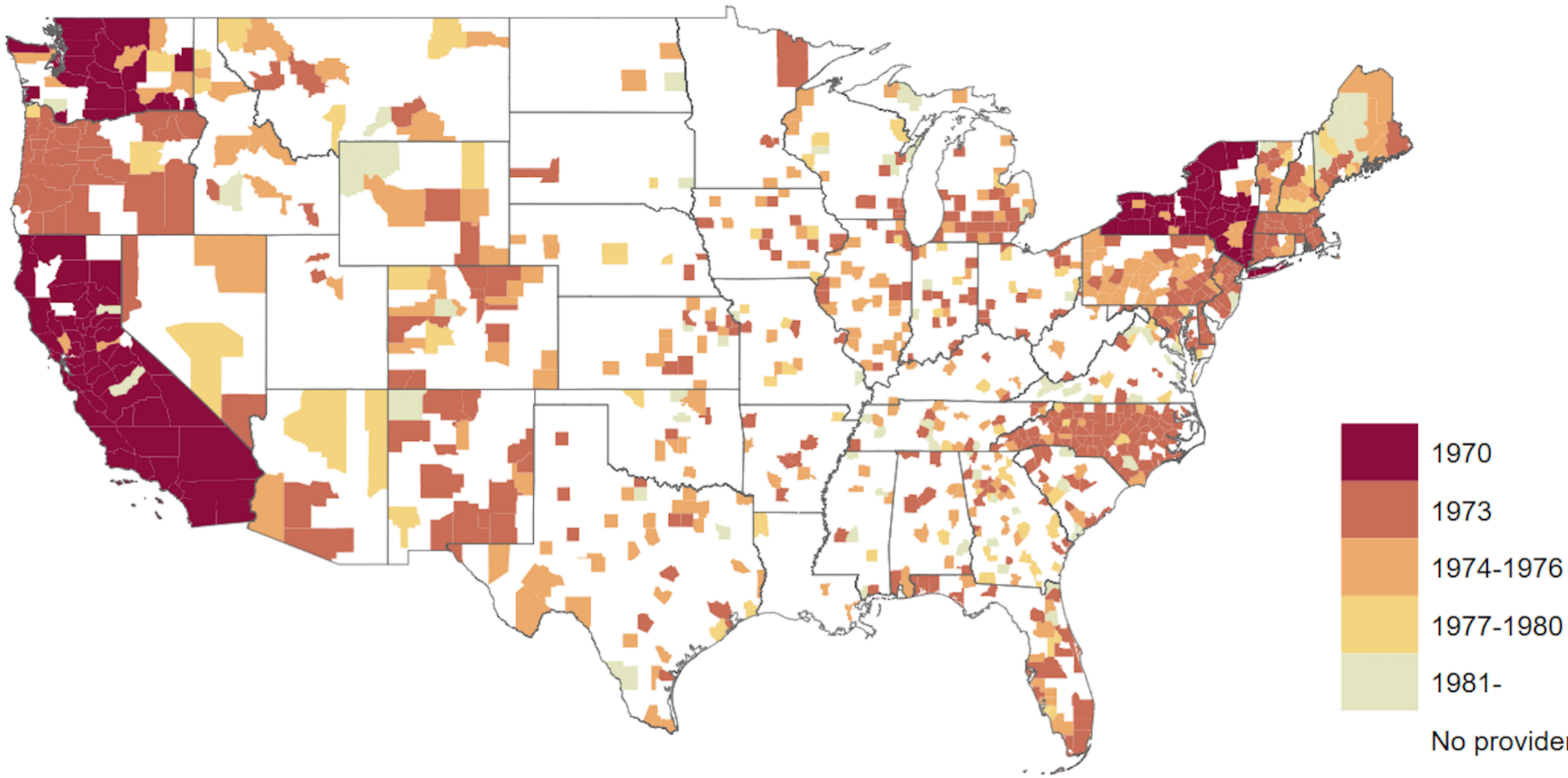


B. First Births

C. Non-First Births

Notes: Point estimates of cohort indicators interacted with the early-legalizing indicator are plotted. Sample includes individuals from PSID Childbirth and Adoption History File whose mother's most recent observation is after she turns 40 years old. Covariates include gender, race, mother's age at birth, and categories of maternal education. Regression controls for year of birth and state of birth fixed effects. Confidence intervals presented are on the 90% level. Heteroskedasticity-robust standard errors clustered at the state level.

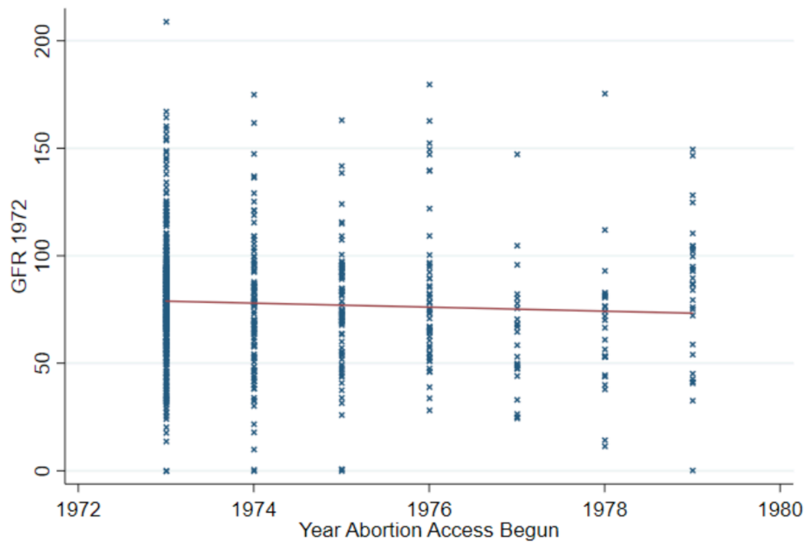
Figure 1.2. Abortion Service Providers Roll-Out, 1970-1987



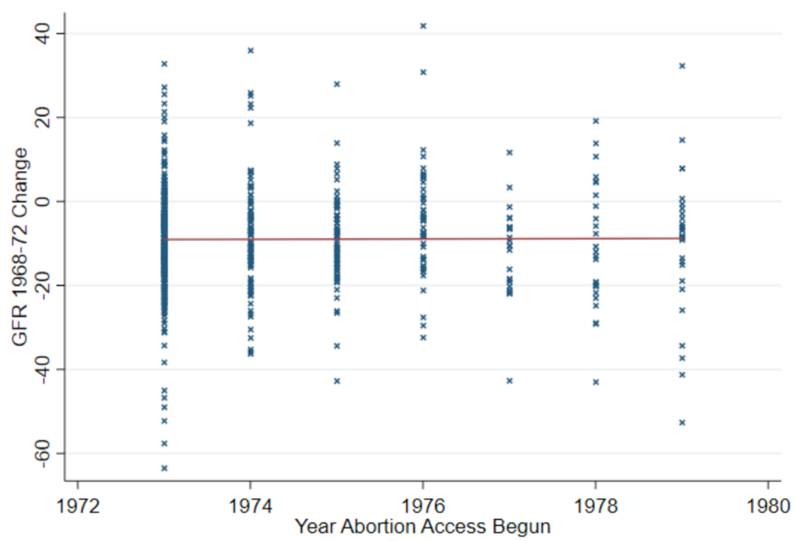
Notes: Number of providers between 1970 to 1972 in early-legalizing states are assumed to be identical to the number observed in 1973.
Source: Guttmacher Institute.

Figure 1.3. County Fertility Rates and Service Roll-Out

A. Year of Roll-Out and General Fertility Rate (GFR)

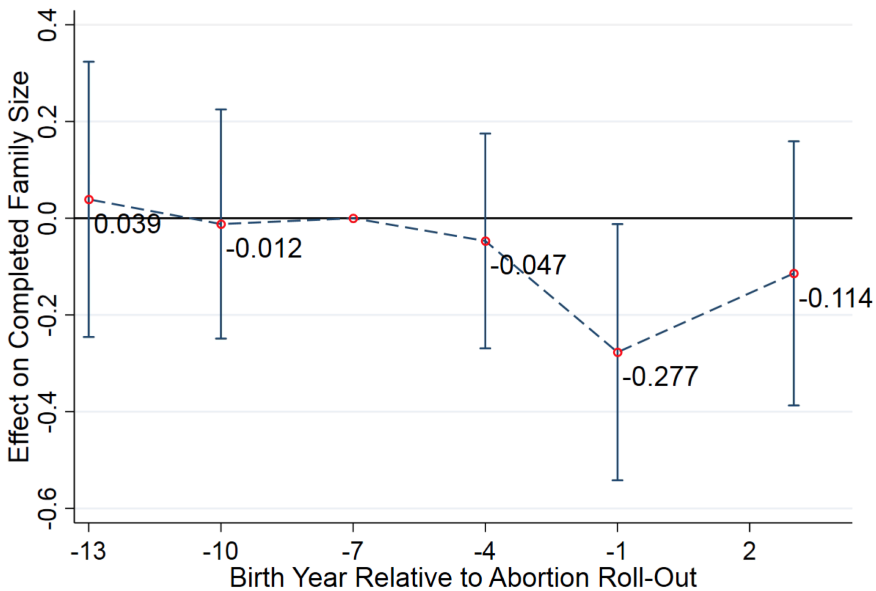


B. Year of Roll-Out and 1968 to 1972 Change in the GFR



Notes: The line indicates the estimated relationship between the GFR or the change in GFR and the year of first abortion service provider in the county. The slope estimate for panel A is -0.93 (standard error 0.67). The slope estimate for panel B is 0.04 (standard error 0.27). *Source:* NCHS Natality Detail Files, 1968 and 1972.

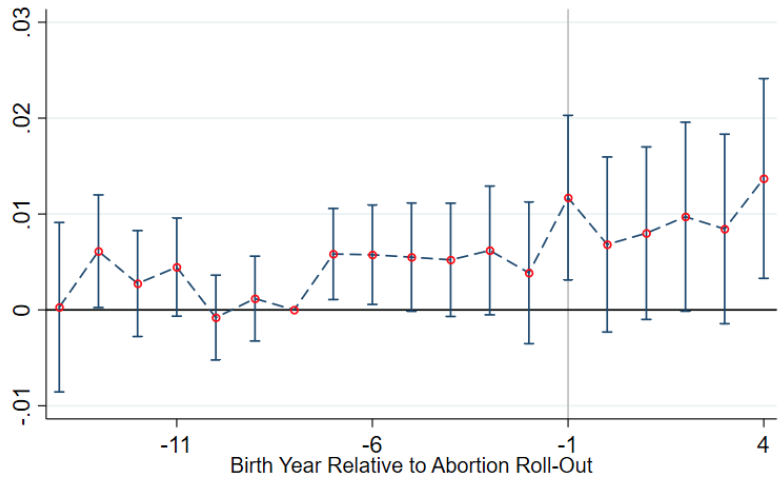
Figure 1.4. Abortion Roll-Out and Family Size, Event-Study Estimates



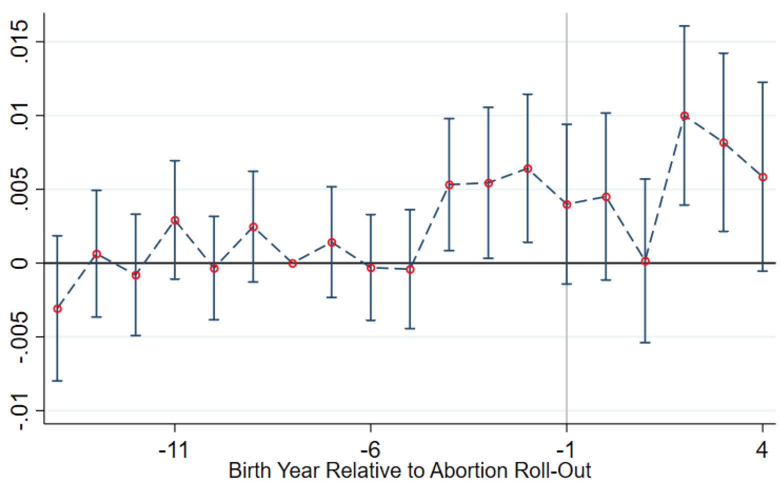
Notes: Dependent variable is completed family size. Point estimates of event-time grouped into three-year bins are plotted. Sample includes individuals from PSID Childbirth and Adoption History File whose mother's most recent observation is after she turns 40 years old. Regression controls for state-by-cohort fixed effects and county fixed effects. 90-percent, point-wise confidence intervals for each estimate are presented. Heteroskedasticity-robust standard errors clustered at the county level. *Source:* Restricted PSID Individual and Childhood and Adoption History File.

Figure 1.5. Effects of Abortion Roll-Out on Long-Run Indices

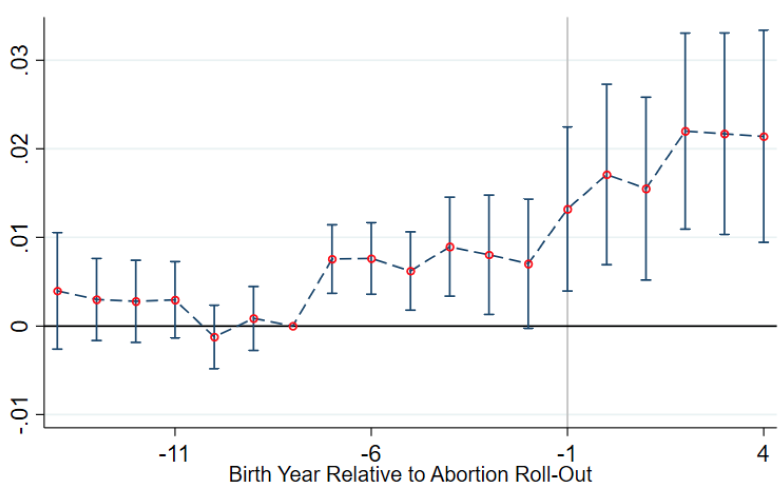
A. Human Capital Index



B. Economic Self-Sufficiency Index

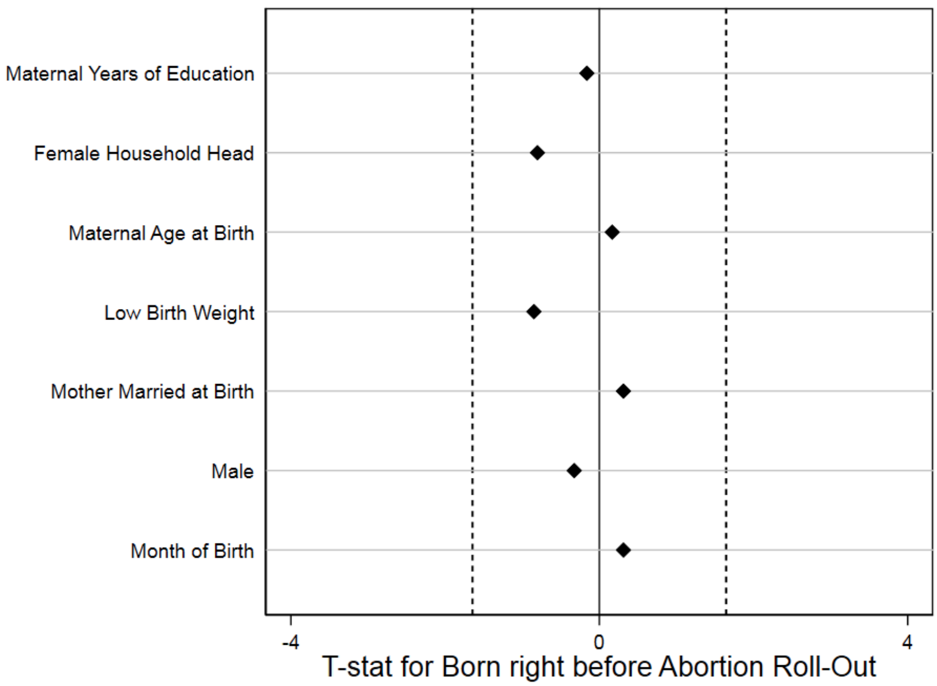


C. Neighborhood Quality Index



Notes: Each figure plots event-study estimates for long-run outcomes using the specification in equation (2). Standard errors clustered at the county level. 90-percent, point-wise confidence intervals for each estimate are presented. *Source:* Restricted Census/ACS and SSA Numident file.

Figure 1.6. Test of Potential Threats to Identification



Notes: Sample includes individuals from PSID Childbirth and Adoption History File whose mother's most recent observation is after she turns 40 years old. Each point plotted is the event-time estimate of the -1 to -3 bin. Heteroskedasticity-robust standard errors clustered at the state level. Vertical dashed lines are drawn at ± 1.645. *Source:* Restricted PSID Individual and Childhood and Adoption History File.

Table 1.1. Year of Abortion Provider Roll-Out and 1970 County Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Independent Variable	Total population, 1970	% population growth 60-70	% pop growth net migration 60-70	% female, 1970	% urban, 1970	% nonwhite, 1970	% of population aged 0-4 years, 1970	% of population aged 65+ years, 1970
Coefficient	-3.16e-07***	0.00879	0.00917	-0.0618	-0.0357***	-0.0162	0.191	-0.00450
(s.e.)	(1.04e-07)	(0.00770)	(0.00810)	(0.0474)	(0.00899)	(0.0111)	(0.147)	(0.0335)
Observations	779	779	779	779	779	779	779	779
R-squared	0.256	0.245	0.244	0.241	0.254	0.244	0.244	0.240

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Independent Variable	Median age (years), 1970	Birth rate per 1,000 population, 1968	Death rate per 1,000 population, 1968	Median years schooling completed (1970)	% less than 5 yrs schooling (1970)	% 12 or more yrs schooling (1970)	Median family income, all families, 1969 ($)	Leading party = Democratic
Coefficient	-0.0209	-0.00333	-0.0368	-0.117	0.0409	-0.00148	-6.80e-05	0.111
(s.e.)	(0.0230)	(0.0435)	(0.0305)	(0.114)	(0.0290)	(0.0101)	(8.01e-05)	(0.194)
Observations	779	779	779	779	779	779	779	779
R-squared	0.240	0.240	0.242	0.241	0.242	0.240	0.241	0.240

Notes: Dependent variable is year of provider roll-out in each county. Each column presents coefficient from a simple linear regression. *Source:* 1970 County and City Data Book.

Table 1.2. Number of Younger Siblings and Service Intensity since Roe v. Wade

	# Younger Siblings		Any Younger Siblings	
	(1)	(2)	(3)	(4)
Born 1964-66	-0.106	-0.136	-0.125**	-0.139**
	[0.231]	[0.228]	[0.054]	[0.053]
Born 1967-69	0.450	0.375	0.035	0.009
	[0.328]	[0.317]	[0.100]	[0.098]
Born 1970-72	0.795**	0.755*	0.073	0.050
	[0.387]	[0.376]	[0.130]	[0.129]
Born 1964-66 x # Prov/1k	0.541	0.712	0.689	0.745
	[2.521]	[2.463]	[0.742]	[0.730]
Born 1967-69 x # Prov/1k	-1.606	-1.538	-0.428	-0.439
	[1.904]	[2.059]	[0.636]	[0.655]
Born 1970-72 x # Prov/1k	-3.312**	-3.175**	-0.430	-0.368
	[1.590]	[1.523]	[0.870]	[0.873]
State FE x Trends	✓	✓	✓	✓
County Covariates x Trends		✓		✓
Observations	1,713	1,713	1,713	1,713
Mean Y	1.300	1.300	0.680	0.680
Mean # Service Prov	0.0700	0.0700	0.0700	0.0700

Notes: Sample includes individuals from PSID Childbirth and Adoption History File born between 1960 and 1972 whose mother's most recent observation is after she turns 40 years old. Covariates include gender, race, mother's age at birth, and categories of maternal education. Heteroskedasticity-robust standard errors clustered at the state level.

Table 1.3. Summary of Event-Study Estimates on Long-Run Outcomes

Dependent Variables *(N=15,891,000)*	(1) Control Mean	(2) Event Study at -1 (s.e.)	(3) F −3 to −1 joint	(4) F left of −8 joint	(5) ATET	(6) ATET %
A. Human capital index		0.0117 (0.00522)	2.835	1.191	0.0422	
Years of schooling	13.790	0.0404 (0.01810)	3.476	1.317	0.1458	1.1%
High school or GED completed	0.932	0.0027 (0.00151)	2.087	1.139	0.0097	1.0%
Completed 4 year college	0.332	0.0073 (0.00336)	2.642	0.935	0.0265	8.0%
Professional or doctoral degree	0.031	0.0025 (0.00081)	3.416	1.812	0.0090	29.1%
Has a professional job	0.370	0.0033 (0.00274)	0.809	0.883	0.0119	3.2%
B. Economic self-sufficiency index		0.0040 (0.00329)	1.598	1.401	0.0144	
Worked last year	0.857	-0.0001 (0.00174)	1.157	0.468	-0.0003	0.0%
Number of weeks worked last year	40.990	-0.0192 (0.09520)	1.925	0.794	-0.0693	-0.2%
Usual hours works per week	35.590	0.1200 (0.09290)	0.919	1.134	0.4332	1.2%
Positive labor income (wage only)	0.810	0.0008 (0.00193)	1.331	0.338	0.0027	0.3%
Log labor income (wage only)	10.610	0.0084 (0.00471)	1.089	1.035	0.0303	
In poverty	0.100	0.0006 (0.00200)	0.851	1.389	0.0021	2.1%
C. Neighborhood quality index		0.0132 (0.00562)	2.153	0.504	0.0477	
Mean home ownership in tract	0.730	0.0034 (0.00181)	1.401	1.334	0.0123	1.7%
Positive small family income in tract	0.941	0.0013 (0.00039)	3.594	1.511	0.0045	0.5%
Log family income to poverty ratio in tract	5.939	0.013 (0.00442)	2.679	0.664	0.0451	
Share single head of family in tract	0.434	-0.001 (0.00093)	0.361	1.153	-0.0026	-0.6%
Share children in poverty in tract	0.201	-0.004 (0.00146)	2.863	1.031	-0.0138	-6.9%

Notes: In Column 1, the control mean is calculated using the individuals born 8 years before abortion access. Column 2 presents the estimated intent-to-treat (ITT) effect evaluated at event-time −1 using the specification in equation (2) (see Figure 1.5). Column 3 and Column 4 present test statistics for joint-significance of event-time −3 to −1 and of event-time smaller than −8, respectively. The average treatment-effect-on-the-treated (ATET) estimate in Column 5 divides the ITT effect by the estimated reduction in completed family size, 0.277. Column 6 computes the percentage change implied by the ATET relative to the control mean (the ratio of Column 5 to Column 1). Percent change is inapplicable and therefore left blank for standardized indices and for logged outcomes. *Source:* Restricted Census/ACS and SSA Numident file.

Table 1.4. Event-Study Estimates on Human Capital and Self-Sufficiency, by Gender

Dependent Variables	(1) Control Mean	(2) Event Study at -1 (s.e.)	(3) F -3 to -1 joint	(4) F left of -8 joint	(5) ATET	(6) ATET %
A. Male (N=7,601,000)						
i. Human capital index		0.0189	3.033	0.692	0.0682	
		(0.00648)				
Years of schooling	13.730	0.0671	3.495	0.977	0.2422	1.8%
		(0.02270)				
High school or GED completed	0.923	0.0041	1.717	0.389	0.0147	1.6%
		(0.00202)				
Completed 4 year college	0.328	0.0122	3.104	0.676	0.0440	13.4%
		(0.00411)				
Professional or doctoral degree	0.036	0.0028	2.513	1.025	0.0100	27.7%
		(0.00119)				
Has a professional job	0.357	0.0086	2.426	0.345	0.0310	8.7%
		(0.00351)				
ii. Economic self-sufficiency index		0.0110	3.817	1.065	0.0397	
		(0.00385)				
Worked last year	0.920	0.0035	1.687	0.970	0.0127	1.4%
		(0.00169)				
Number of weeks worked last year	45.110	0.1870	2.845	1.111	0.6751	1.5%
		(0.09800)				
Usual hours works per week	41.460	0.3100	3.776	0.709	1.1191	2.7%
		(0.10300)				
Positive labor income (wage only)	0.865	0.0046	2.047	1.490	0.0165	1.9%
		(0.00200)				
Log labor income (wage only)	10.880	0.0124	2.167	2.384	0.0448	
		(0.00552)				
In poverty	0.074	-0.0036	1.994	1.424	-0.0130	-17.6%
		(0.00191)				
B. Female (N=8,290,000)						
i. Human capital index		0.004	1.175	0.989	0.0149	
		(0.00551)				
Years of schooling	13.850	0.014	1.960	1.059	0.0513	0.4%
		(0.01880)				
High school or GED completed	0.940	0.001	1.154	1.155	0.0052	0.6%
		(0.00165)				
Completed 4 year college	0.337	0.003	1.328	0.836	0.0094	2.8%
		(0.00372)				
Professional or doctoral degree	0.026	0.002	2.293	1.950	0.0085	32.3%
		(0.00097)				
Has a professional job	0.384	-0.002	0.168	0.728	-0.0083	-2.2%
		(0.00336)				
ii. Economic self-sufficiency index		-0.001	1.538	1.178	-0.0034	
		(0.00390)				
Worked last year	0.797	-0.003	1.589	1.016	-0.0117	-1.5%
		(0.00247)				
Number of weeks worked last year	37.160	-0.197	2.126	0.760	-0.7112	-1.9%
		(0.13200)				
Usual hours works per week	30.130	-0.048	0.358	1.432	-0.1736	-0.6%
		(0.11200)				
Positive labor income (wage only)	0.759	-0.003	1.599	0.886	-0.0098	-1.3%
		(0.00266)				
Log labor income (wage only)	10.320	0.001	0.139	0.735	0.0035	
		(0.00623)				
In poverty	0.123	0.004	1.845	1.283	0.0145	11.8%
		(0.00258)				

Notes: See Table 3 notes.

Table 1.5. Event-Study Estimates on Neighborhood Quality, by Gender

Dependent Variables	(1) Control Mean	(2) Event Study at -1 (s.e.)	(3) F -3 to -1 joint	(4) F left of -8 joint	(5) ATET	(6) ATET %
A. Male (N=7,601,000)						
Neighborhood quality index		0.0168	2.708	1.075	0.0606	
		(0.00599)				
Mean home ownership in tract	0.745	0.0026	1.115	2.339	0.0094	1.3%
		(0.00196)				
Positive small family income in tract	0.943	0.0015	4.276	1.155	0.0054	0.6%
		(0.00045)				
Log family income to poverty ratio in tract	5.941	0.0145	3.399	1.782	0.0523	
		(0.00471)				
Share single head of family in tract	0.434	-0.0012	1.178	2.197	-0.0043	-1.0%
		(0.00105)				
Share children in poverty in tract	0.200	-0.0044	2.830	0.928	-0.0160	-8.0%
		(0.00158)				
B. Female (N=8,290,000)						
Neighborhood quality index		0.010	1.509	0.664	0.0347	
		(0.00605)				
Mean home ownership in tract	0.749	0.004	2.798	0.865	0.0147	2.0%
		(0.00201)				
Positive small family income in tract	0.942	0.001	2.624	1.226	0.0038	0.4%
		(0.00039)				
Log family income to poverty ratio in tract	5.938	0.010	2.032	0.663	0.0372	
		(0.00467)				
Share single head of family in tract	0.433	0.000	1.830	0.650	-0.0006	-0.2%
		(0.00107)				
Share children in poverty in tract	0.202	-0.003	2.103	0.665	-0.0116	-5.7%
		(0.00152)				

Notes: See Table 3 notes.

Table 1.6. Effects Heterogeneity Interaction between Family Size and Public Programs

	Abortion Access Interacted with:			
	Head Start	Food Stamps	Medicaid Expenditure	Community Action Programs
Outcomes:				
Human Capital Index	0.0157	0.00551	0.00805	-0.00617
(s.e.)	(0.00612)	(0.00486)	(0.00857)	(0.00456)
Economic Self-Sufficiency Index	0.00467	0.00579	-0.00301	-0.00327
(s.e.)	(0.00401)	(0.00316)	(0.00548)	(0.00337)
Neighborhood Quality Index	0.0098	0.00137	0.00997	-0.00838
(s.e.)	(0.00394)	(0.00409)	(0.00822)	(0.00356)

Notes: The empirical model is similar to equation (2) but additionally includes each event-time dummy interacted with an indicator of high exposure to a public program, $\sum_y \psi_y D_c 1\{b - T_c = y\} Abortion_c$. Estimates of the interaction effect ψ_y are presented. Estimates are intent-to-treat (ITT) effects evaluated at event-time -1. *Source:* Restricted Census/ACS and SSA Numident file.

CHAPTER II

Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency

Abstract

This paper evaluates the long-run effects of Head Start using large-scale, restricted 2000-2013 Census-ACS data linked to date and place of birth in the Social Security Administration's Numident file. Using the county-level roll-out of Head Start between 1965 and 1980 and state age-eligibility cutoffs for school entry, we find that childhood participation in Head Start is associated with increases in adult human capital and economic self-sufficiency, including a 0.29-year increase in schooling, a 2.1-percent increase in high-school completion, an 8.7-percent increase in college enrollment, and a 19-percent increase in college completion. These estimates imply sizable, long- term returns to investing in large-scale preschool programs.

JEL Codes: I2, J24, J6

Convincing evidence on the longer-term impacts of scaled-up pre-k programs on academic outcomes and school progress is sparse, precluding broad conclusions. The evidence that does exist often shows that pre-k-induced improvements in learning are detectable during elementary school, but studies also reveal null or negative longer-term impacts for some programs.

– Brookings Pre-Kindergarten Task Force of Interdisciplinary Scientists, (Phillips et al., 2017)

2.1 Introduction

In 1965, the U.S. began a new experiment in the provision of public preschool for disadvantaged children. The motivation was simple: "the creation of and assistance to preschool, day care, or nursery centers for 3- to 5-year-olds will provide an opportunity for a head start by canceling out deficiencies associated with poverty that are instrumental in school failure" (United States Senate Committee on Labor and Public Welfare, 1964). The program that ensued is the now-famous "Head Start," a "prep school for poor kids" which aimed to help millions of children escape poverty (Levitan, 1969).

More than fifty years later, Head Start is one of the most popular of the War on Poverty's programs, serving around 900,000 children annually at a cost of \$9.6 billion in 2017 to the federal government. Unlike expensive, small-scale "model" programs such as Perry Preschool and Abecedarian, Head Start's architects prioritized widespread access, calculating that a massive preschool expansion would maximize its poverty-fighting (and political) benefits. Skepticism about the quality of this large-scale preschool program coupled with difficulties in evaluation have generated controversy about its short-term benefits for decades (Duncan and Magnuson, 2013; Currie, 2001; Westinghouse Learning Corporation, 1969). Convincing evidence regarding Head Start's long-term effects has remained even more elusive, thanks to the lack of program randomization in its early years, small sample sizes of longitudinal surveys, and the difficulty of measuring adults' access to Head Start decades ago. Consequently, the best estimates of Head Start's long-term effects are limited by lingering concerns about endogeneity (sibling comparison designs, Currie and Thomas (1995); Garces, Thomas and Currie (2002); Deming (2009) and imprecision (measurement error in funding and access, Ludwig and Miller (2007); Carneiro and Ginja (2014). Whether Head Start achieved its goal of increasing life opportunities for children remains an open question. This paper uses large-scale data to estimate Head Start's long-term effects on human capital and economic self-sufficiency. By linking the restricted long-form 2000 Census and 2001-2013 American Community Surveys (ACS) to the exact date and place of birth from the Social Security Administration's (SSA) Numident file, we observe outcomes for one-quarter of U.S. adults as well as a high quality measure of their access to and eligibility for Head Start as children. The resulting sample is four orders of magnitude larger than longitudinal surveys, and information on place of birth and exact date of birth ameliorates (potentially non-classical) measurement error in childhood access to Head Start.

Our research design exploits the county-level roll-out of Head Start programs from 1965 to 1980 at the Office of Economic Opportunity (OEO) (Levine, 1970; Bailey and Goodman-Bacon, 2015; Bailey and Duquette, 2014; Bailey and Danziger, 2013; Bailey, 2012). This approach exploits the well-documented "great administrative confusion" at the OEO (Levine, 1970), mitigating problems of measurement error in archival funding data (Barr and Gibbs, 2017) and concerns about the endogeneity of Head Start funding levels. An additional strength of our design is that it leverages Head Start's age-eligibility guidelines, comparing cohorts who were age-eligible when it launched (ages 5 and younger) to cohorts born in the same county that were age-ineligible (children 6 and older). The substantial variation in adult outcomes means that even our large dataset does not permit the estimation of a regression discontinuity or regression kink design, but our approach is based on similar assumptions. Much like a regression kink design (Card et al., 2015), our key identifying assumption is that

Head Start's causal effect is the only reason for a change in the relationship between a child's age at the program's launch and her outcomes as an adult. Because even our large-scale data are not large enough to estimate a regression kink formally, we present both event-study estimates as well as trend-break tests based on spline parameterizations.

The results suggest that Head Start increased the human capital and economic self-sufficiency of disadvantaged children. An index of adult human capital rose by 10 percent of a standard deviation among Head Start participants relative to children born in the same county who were age 6 when the program began. Participating children achieved 0.29 more years of education, were 2.1 percent more likely to complete high school, 8.7 percent more likely to enroll in college, and 19 percent more likely to complete college. In addition, Head Start increased economic self-sufficiency in adulthood by almost 4 percent of a standard deviation – gains driven largely by a 12-percent reduction in adult poverty and a 29-percent reduction in public assistance receipt. We find no evidence of reductions in incarceration. Heterogeneity in Head Start's effects suggests that they reflect, in part, practices outside of pre-school curriculum. In particular, health screenings and referrals as well as more nutritious meals appear to be important mechanisms for the program's effects on disadvantaged children. In addition, the effects of Head Start appear to be complement greater family and public resources arising from a stronger economy. Overall, Head Start appears to have achieved the goals of its early architects, both increasing children's economic opportunities and reducing poverty.

A final analysis quantifies the private, internal rates of return to dollars spent on Head Start in the 1960s and 1970s. Rather than using changes in wage income directly, we use the National Longitudinal Survey of Youth 1979 (NLSY79) to predict changes in earnings for the relevant cohorts net of any ability differences (Neal and Johnson, 1996; Deming, 2009). Using potential earnings accounts for Head Start-induced negative selection in men's employment (driven by reductions in disability) and positive selection in women's employment (due to the income effect dominating for the least skilled women). This exercise suggests a private internal rate of return to Head Start of 7.7 percent, which ranges from around 4 percent for women to 11 percent for men. Using only savings on public assistance expenditures as a conservative method to calculate the program's social returns, we find that the internal rate of return of putting one child through Head Start is 2.4 percent. In short, this paper's estimates suggest substantial long-run returns to America's first scaled-up, public preschool program. While the results do not imply that all of today's large-scale preschool programs work, they suggest that some less-than-model preschool programs may have lasting effects – a key finding for current policy deliberations (Phillips et al., 2017).

2.2 The Launch of Head Start in the 1960s and Expected Effects

In the 1960s, the idea that preschool could improve children's cognitive development was revolutionary. Challenging the conventional notion that IQ was immutable and fixed at birth, Joseph McVicker Hunt's (1961) book, Intelligence and Experiences, persuasively argued that children's intelligence could be significantly improved by altering their experiences. Benjamin Bloom further emphasized that the first four years of children's lives was a "critical period," noting that "intelligence appears to develop as much from conception to age 4 as it does from age 4 to 18" (Bloom, 1964). This idea suggested an innovative strategy for poverty prevention. Because poor children started school with significantly less educational background, comprehensive preschool could give them a "Head Start," improving their success in school and addressing a root cause of poverty.

2.2.1 A Brief History of Head Start's Launch

Funded by the OEO, Head Start began as an 8-week summer program in 1965. After a successful first summer, President Lyndon Johnson announced that Head Start would become a full-year program for children ages 3 to 5. The director of the OEO wrote 35,000 letters to public health directors, school superintendents, mayors and social services commissioners to encourage applications. The OEO also made a special effort to generate applications in America's 300 poorest counties (Ludwig and Miller, 2007).

Head Start's political popularity led to an even faster launch than other War on Poverty programs. Figure 2.1 shows the program's quick expansion. By 1966, Head Start had begun in more than 500 counties where over half of the nation's children under age 6 resided. By 1970, federal expenditures on the program reached $326 million, or $2.1 billion in 2018 dollars (OEO, 1970). This early expansion ensured that by 1970, Head Start existed in roughly half of U.S. counties, putting preschool programs within a short drive for 83 percent of children under age six (Appendix Table B.1).

The exact timing of Head Start's launch depended on many idiosyncratic factors. The OEO's "wild sort of grant-making operation" has been well documented in oral histories (Gillette 1996: 193) as well as in more recent, quantitative analyses (Bailey and Duquette, 2014). In the case of Head Start, other factors were key as well: how excited were local institutions or politicians about the program? Was their there adequate and available space to launch? Could the program be integrated within the public school system or would it remain separate? The final result of the grant- making process and local constraints was that Head Start began earlier in areas that were significantly more populous and urban (Appendix Table B.2) – areas where more children could be served. In addition, urban areas were funded earlier. After accounting for population and urban differences, the roll-out of Head Start was not strongly affected by other county characteristics (Appendix Table B.3).

Consistent with the historical evidence that this national program was rushed into existence, exactly when Head Start began after 1965 does not appear to be systematically related to pre-existing local characteristics.

2.2.2 Head Start's Mission

Head Start's architects adopted a holistic approach that aimed to develop children's mental and physical abilities by improving health; self-confidence; verbal, conceptual, and relational skills; and raising parental involvement. Levitan (1969) notes that Head Start's 1966-7 budget included early childhood education (daily activities and transport, 70 percent), health services (including immunizations, screenings and medical referrals, 4 percent), and nutrition (14 percent). Parent involvement, social services (e.g., helping families cope with crises), and mental health services accounted for the remaining budget.

The expected effects of the program on adult outcomes could flow directly from the early learning facilitated by the program. But the role of health services and nutrition may be important as well. Head Start's vaccinations and screening (e.g., tuberculosis, diabetes, vision, hearing) and referrals to local physicians may have prevented complications from childhood diseases (North, 1979; Ludwig and Miller, 2007) and helped parents obtain simple, cost-effective technologies to improve learning (e.g., eye glasses and hearing aids or antibiotics to reduce hearing damage from ear infections). Healthy meals and snacks may have also raised children's ability to learn. Early estimates suggest that more than 40 percent of children entering Head Start were receiving less than two-thirds of the recommended allotment of iron, and 10 percent were extremely deprived in terms of their daily calories (Fosburg et al., 1984). Among children who received blood tests in the 1968 full-year program, 15 percent were found to be anemic (Office of Child Development , DHEW). Reducing these nutritional deficiencies could also translate into significant gains in educational achievement in both the short and longer term (Frisvold, 2015).

The challenges of quickly starting a new national program meant that implementation often deviated from ideals. Not only did Head Start lack curricular standardization, but programs struggled to find high-quality teachers to achieve the suggested pupil-to-teacher ratio of 15:1. As a practical solution many centers relied on para-professionals, most of whom lacked post-secondary education; thirty percent had not finished high school (Hechinger, 1966; Braun and Edwards, 1972). In addition, many components of Head Start phased in slowly. For instance, the OEO wrote that in 1965, "the proportion of children receiving treatment for conditions discovered in Head Start medical and dental examinations ... was probably under 20 percent. It rose to over 65 percent in 1966, and in 1967 we fully expect it to have reached over 90 percent" (OEO, 1967).

Consequently, Head Start in its earliest years was far from a model preschool program. Nevertheless, even the less-than-ideal implementation of Head Start was likely higher quality than the alternatives available to low income children in the 1960s (Currie, 2001). Importantly, similar concerns hold today: Head Start's quality score from the National Institute for Early Education Research places Head Start program quality around the median of the score distribution (Espinosa, 2002) but the program may still be much better than informal child care (Loeb, 2016).

2.3 Literature Regarding the Long-Term Effects of Head Start

Previous evaluations of Head Start provide suggestive evidence of the program's long-term effects on human capital and economic self-sufficiency. One pioneering approach was the use of family fixed effects with longitudinal data. Building on work by Currie and Thomas (1995), Garces, Thomas and Currie (2002) used the Panel Study of Income Dynamics (PSID) to compare children who participated in Head Start to their siblings who did not. They show that Head Start increased high school graduation rates and college enrollment among whites and reduced arrest rates among blacks. Using a similar research design for more recent cohorts in the National Longitudinal Survey of Youth (NLSY), Deming (2009) finds that Head Start participation had large and positive effects on a summary index of adult outcomes (including high school graduation, college attendance, "idleness," crime, teen parenthood, and health status). Well-known critiques caution that sibling comparisons may suffer from sources of endogeneity bias (Griliches, 1979; Bound and Solon, 1999). In addition, small sample sizes in longitudinal surveys may provide unreliable estimates of Head Start's effects (Grosz, Miller and Shenhav, 2017).

More recent work exploits shifts in access to Head Start using three distinct research designs. The path-breaking application of RD in Ludwig and Miller (2007) exploited the OEO's special effort to generate grant proposals from the 300 poorest counties. Comparing the outcomes of children on either side of this threshold, they find evidence that Head Start reduced childhood mortality and increased the receipt of high-school degrees and college enrollment. However, because the 1990 and 2000 Censuses required them to use county of residence in adulthood to proxy for childhood Head Start access, measurement error causes their education results to be sensitive to specification and often statistically insignificant. Carneiro and Ginja (2014) use an RD in state-, year-, and household-based income eligibility cutoffs for more recent Head Start programs. They find that Head Start decreased behavioral problems, the prevalence of some health conditions (including obesity) between the ages of 12 and 17, and crime rates around age 20. They find a positive though statistically insignificant effect on receiving a high-school diploma as well as suggestive evidence that Head Start reduced college enrollment.

In work closely related to this paper, three studies make use of county-year variation in Head Start funding in the 1960s and 1970s to quantify the program's long-term effects. Using a sample of likely eligible children from the NLSY, Thompson (2017) finds that greater funding for Head Start at ages 3 to 6 raised college graduation rates, reduced the incidence of health limitations, and tended to raise adult household income. Focusing on a "high impact" sample, Johnson and Jackson (2017) find that an average level of Head Start and education spending increases the likelihood that children graduated from high school by 8 p ercentage points and gained 0.39 years of schooling. These children also experienced a 7.8 log-point increase in adult wages, a 14.4 log-point increase in adult family income at ages 20 to 50, a 3.6 percentage-point reduction in poverty at ages 20 to 50, and a 3 percentage-point reduction in adult incarceration. Finally, Barr and Gibbs (2017) examine the intergenerational effects of Head Start using the NLSY and two research designs: family fixed effects and variation in program availability across birth counties (also referred to as "roll-out"). To alleviate concerns about the endogeneity of funding levels and measurement error in the National Archives data, their roll-out design uses a binary measure of Head Start access that is equal to one if funding exceeds the 10th percentile of observed funding per four-year-old. They find evidence of large first-generation effects on women (including a gain of a half a year of schooling) and large second-generation effects on their children's high school graduation and completed education.

2.4 Data and Research Design

This study combines the long-form 2000 Census and 2001-2013 ACS with the SSA's Numident file to shed new light on Head Start's long-term effects. The Census/ACS data represent almost one quarter of the U.S. population and are four orders of magnitude larger than previously used longitudinal samples. Another advantage of these combined data is that the Numident contains county of birth (rather than adulthood residence) and exact date of birth, which allows a high-quality proxy for Head Start access and age eligibility in childhood. The data's main disadvantage is that they contain no information on family background. This lack of covariates means that we cannot model many determinants of adult outcomes – which limits precision even in this large dataset – or model treatment effect heterogeneity by childhood characteristics.

Our sample is comprised of children born from 1950 to 1980 in U.S. states where the school-entry age cutoff is known. We additionally limit our sample to individuals who are in their prime earning years (ages 25 to 54). We collapse these data to means by birth year, survey year, county of birth, and school age. We also weight our regressions using the number of observations in each cell (Solon, Haider and Wooldridge, 2015). To minimize disclosure concerns at the Census Bureau, we use only observations with non-allocated and non-missing values for all outcomes.

Our outcomes of interest are summary measures of human capital and economic self-sufficiency, which permit tests of co-movements of related adult outcomes and limit the number of statistical tests (Kling, Liebman and Katz, 2007). A shortcoming of this approach is that, because indices weight each component equally, large changes in one dimension are averaged with potentially opposite-signed or zero effects in other dimensions. We, therefore, also examine the individual index components. The human capital index includes four binary variables indicating achievement of a given level of education or greater: high school or GED, some college, a 4-year college degree, and a professional or doctoral degree; years of schooling, and an indicator for working in a professional occupation. Our index of self-sufficiency includes binary indicators of employment, poverty status, income from public sources, family income, and income from other non-governmental sources; continuous measures of weeks worked, usual hours worked, the log of labor income, log of other income from non-governmental sources, and log ratio of family income to the federal poverty threshold.

2.4.1 Measuring Exposure to Head Start

Combining data on the launch of Head Start programs from Bailey and Goodman-Bacon (2015) with the Census/ACS-Numident permits two refinements to previously used research designs Barr and Gibbs (2017); Johnson and Jackson (2017); Thompson (2017). First, we use only variation in the launch of the Head Start program rather than a continuous measure of Head Start spending. This refinement (1) addresses the potential endogeneity of Head Start funding levels to the program's performance and (2) sidesteps issues of measurement error in the National Archives grant data (Barr and Gibbs, 2017). Second, we examine changes in outcomes for children who were age-eligible for Head Start (ages 3-5 or younger) relative to those who were age-ineligible (ages 6+) when it launched, allowing for the effects to vary by the number of years each cohort was potentially eligible. Age eligibility is based on exact date of birth in the Numident and school-entry age cutoffs, which alleviates measurement error in defining the potential treatment and control groups. Finally, our large dataset allows us to use state-by-birth-year fixed effects to adjust estimates for state economic and policy changes that could have affected children's outcomes independently of Head Start. Our identifying assumption in the analysis that follows is that the causal effect of Head Start is the only reason for a change in the relationship between a child's age at the program's launch and her outcomes as an adult. (See Appendix for more description of our sample.)

2.4.2 Event-Study Regression Model

Our research design uses a flexible event-study framework and roll-out of Head Start to estimate the effect of exposure to Head Start on long-term human capital and economic outcomes,

$$Y_{bct} = \theta_c + \alpha_t + \delta_{s(c)b} + \boldsymbol{Z_c'} b \boldsymbol{\beta} + HeadStart_c \left[\boldsymbol{Age'_{bs(c)} \phi} \right] + \epsilon_{bct} \tag{2.1}$$

Children's birth years are indexed by $b = 1950 - 1980$, county of birth by c, and Census/ACS year by $t = 2000 - 2013$. Specifications include fixed effects for county of birth, θ_c, year, α_t, and state-by-birth-year, $\delta_{s(c)b}$, which, respectively, capture time-invariant differences across counties, national changes affecting all individuals, and changes in state policies that differentially affect birth cohorts. Although covariates matter little, we follow the literature and include county characteristics, Z_c interacted with a linear trend in year of birth, b (Hoynes, Page and Stevens, 2011; Bailey, 2012; Bailey and Goodman-Bacon, 2015).

$HeadStart_c$ is a binary variable equal to 1 if a child was born in a county that received a Head Start grant before 1980. Age is a set of dummy variables for a child's "school age" at the time of Head Start's launch, $1(T_c^* - b = a)$ where $a = -15$ to 30 (or $T_c^* = 1965$ and $b = 1980$ to $T_c^* = 1980$ and $b = 1950$) and T_c^* is the year Head Start began in county c. We omit school age 6 (age 6 before the school entry cut-off date), because these children would have been unlikely to have attended Head Start rather than public school. Our point estimates of interest, ϕ, describe the evolution of the intent-to-treat (ITT) effects of Head Start on long-term human capital and economic self-sufficiency. Standard errors are corrected for heteroskedasticity and adjusted for an arbitrary within-birth-county covariance structure (Arellano, 1987; Bertrand, Duflo and Mullainathan, 2004). In our tables, we also report p-values corrected for multiple hypothesis testing using the Bonferroni-Holm method in our tables (Holm, 1979; Duflo, Glennerster and Kremer, 2007).

2.4.3 Expected Effects of Exposure to Head Start by Age at Launch

The event-study model provides a flexible approach that imposes few restrictions on the relationship between Head Start and adult outcomes. Although economic theory does not make predictions as to the magnitudes of the event-study coefficients, the program's phased implementation and the greater potential for some children to enroll (due to multiple years of exposure) predict a specific pattern. Figure 2.2 plots this pattern under the following set of assumptions.

First, if we assume Head Start had no effect on children who were over age 5 when the program launched, then the relationship between adult outcomes and Head Start for these children should be zero. This is the equivalent to a test for a pre-trend in our analysis and is illustrated as a flat line for children ages 6 to 12 in Figure 2.2A.

Second, if Head Start has a positive causal effect on adult outcomes, we expect to see a change in those outcomes for children under age 5 when it launched, because these cohorts would have been the first to have been age-eligible and have access. This would not result in an immediate shift in the level of outcomes (akin to a regression discontinuity, RD) but rather a shift in the slope (akin to a regression kink, RK). The reason is that Head Start's capacity grew over time, both because new programs were added but also as individual programs matured. Program quality also increased over time with better hiring and training of teachers, curriculum development, and the implementation of auxiliary services (e.g., health). Studies of other War on Poverty programs such as family planning or community health centers suggest that many of these programs reached maturity around 4 to 5 years after launch (Bailey, 2012; Bailey and Goodman-Bacon, 2015). Figure 2.2A illustrates this as the implementation curve (line with square markers), which rises from zero to 100 percent.

In addition to these gradual changes in program quality and capacity, we also expect larger effects for children who were younger when Head Start launched, simply because they would have been age-eligible for a larger share of their preschool years. For instance, a child 5 years old when Head Start launched could participate for at most one year, whereas a 3-year-old child would be age-eligible for three years. This does not mean that the 3-year-old enrolled for more than one year. However, it is more likely that a child enrolls if s/he had three years to do so. Figure 2.2A illustrates this cumulative potential access to Head Start as a linear relationship (dashed line with circle markers), but differences in the likelihood of enrollment by age could make this relationship more S-shaped as well (because enrollment in the early years was more likely at ages 4 than 5).

The combination of phased implementation and cumulative potential access to Head Start implies a non-linear change in the relationship between age at Head Start launch and adult outcomes (solid, bold line). In Figure 2.2A's stylized example, children ages -1 or younger at Head Start's launch would have been age-eligible for a fully implemented program for each of their three years of eligibility. Assuming that Head Start did not continue to mature and that it did not have any complementarities with other War on Poverty programs, the relationship should level off for children ages -1 or younger at launch, because all cohorts born after this would have had the same potential exposure to Head Start as the -1 cohort.

Note, however, that Figure 2.2B shows how relaxing two assumptions implies a slightly different shape. First, allowing for effects on children ages 6 and older implies that the curve would begin to slope up before age 6. This is possible because 10 percent of children in full-year Head Start were 6 or older (Vinovskis, 2005), and age-ineligible children could still benefit from their younger siblings' participation (Garces, Thomas and Currie, 2002). Because our subsequent analysis standardizes the effects at age 6 to zero, this relationship would appear as the flat part of the line falling below zero. Second, if the Head Start program continued to mature after 5 years or was complemented by other programs (e.g., Goodman-Bacon (2018) notes that Medicaid continued to expand into the 1970s), we would expect to see a slope for cohorts ages -1 and younger when the program launched.

2.4.4 Spline Summary Specification

Our event-study estimates impose none of the restrictions used to outline the expected effects of Head Start in Figure 2.2. However, we expect estimates from this flexible event-study specification to be noisy, in part because so many unobserved factors determine adult outcomes. To improve precision and test formally for trend breaks, we use Figure 2.2's predictions to guide the specification of a three-part spline with knots at ages 6 and -1. We implement this by replacing the $Age'_{bs(c)}$ in equation 2.1 with components of the spline in age, $a = T_c^* - b$:

$$Y_{bct} = \theta_c + \alpha_t + \delta_{s(c)b} + \boldsymbol{Z_c'}b\boldsymbol{\beta} + HeadStart_c\left[\boldsymbol{D'_{cb}\rho_1} + a\boldsymbol{D'_{cb}\rho_2}\right] + \epsilon_{bct} \tag{2.2}$$

where D'_{cb} is a vector of dummy variables, $1(-10 \leq a \leq -1)$, $1(-1 \leq a \leq 6)$, $1(6 \leq a \leq 15)$, $1(a < -10)$, and $1(15 > a)$ and the other variables remain as previously defined. We constrain the estimates of ρ_1 and ρ_2 to ensure that the spline joins at $a = 6$ and -1. While the spline specification is more restrictive than our flexible event-study, it improves precision by imposing restrictions that should be true. The spline allows a parsimonious method to test for a pre-trend (captured in the slope of the segment for $1(6 \leq a \leq 15)$) as well as a formal trend-break test between components $1(6 \leq a \leq 15)$ and $1(-1 \leq a \leq 6)$. This final test captures whether the relationship between adult outcomes and age at Head Start's introductions changes at age 6 – the age at which older children tend to stop participating in Head Start and begin first grade.

2.5 Tests of Identifying Assumptions

The research design outlined in the previous section relies on two crucial assumptions: (1) the launch of a Head Start program increased participation in Head Start and (2) the launch of a Head Start program did not coincide with other county-level changes that would affect the outcomes of preschool children. This section provides further evidence on both assumptions.

2.5.1 How Much Did Head Start Increase Preschool Enrollment?

While there is little doubt that introducing a Head Start program increased children's attendance in this program, the magnitude of this relationship net of crowd-out is crucial for interpreting the ITT effects recovered in equations 2.1 and 2.2. Administrative data suggest that the launch of a Head Start program significantly increased children's enrollment. The OEO reported that full-year Head Start served over 600,000 children before 1968, rising from 20,000 children in 1965, to 160,000 in 1966, to around 215,000 in 1967 and 1968 (OEO 1965, 1966, 1967, 1968, 1970). About 257,700 children attended full-year Head Start in 1970. Three-quarters of the children were aged 4 or 5, three-quarters were nonwhite, and 62 percent came from families with less than $4,000 in annual income. Between 1971 and 1978, enrollment and directory information suggest that the average county with a Head Start program served roughly 309 children. These sources imply that the average Head Start program served from about 10 percent of resident age-eligible children in 1971 to 15.8 percent in 1978.

If Head Start substituted for private preschool for some children (Cascio and Schanzenbach, 2013; Kline and Walters, 2016; Bassok, Fitzpatrick and Loeb, 2014), administrative data may overstate the role of Head Start programs in increasing exposure to preschool. To examine this possibility, we use the 1970 Census, which was the first to ask children younger than age 5 about school enrollment as of February 1 – a date during the school year, which should capture enrollment in full-year Head Start. The Census data show that four-year-old children in counties without Head Start programs were 3.4 percentage points less likely to be enrolled in school (16.8 versus 20.2 percentage points, see Figure B.1). Five-year-old children were 17 percentage points less likely to be enrolled in school (48.9 versus 65.9 percentage points). These gaps are 5.9 percentage points among 4 year olds and 21.3 percentage points among 5 year olds when looking only at children of mothers with less than a high school education.

We use a linear probability model to adjust these gaps for state fixed effects (to account for age-invariant, state-level factors that determine the local supply of preschools) and 1960 county characteristics (share of county population in urban areas, in rural areas, under 5 years of age, 65 or older, nonwhite, with 12 or more years of education, with less than 4 years of education, in households with income less than $3,000, in households with incomes greater than $10,000, local government expenditures, income per capita, and whether the county was among the 300 poorest counties). Details are presented in Appendix B.6 and summarized here for parsimony. The regression results show that school enrollment was 14.9 percentage points higher for all five year olds, 15.1 percentage points higher for boys, and 14.5 percentage points higher for girls (Appendix Table B.5). These results are highly robust to the inclusion (or exclusion) of different covariates.

Consistent with crowd-out being minimal, the 14.9 percent increase in enrollment in the Census is only slightly smaller than the 15.8 percent contained in administrative data. This estimate is also comparable to other studies. Garces, Thomas and Currie (2002) estimates the national Head Start participation rate was between 10 percent and 17 percent for the 1964 to 1970 cohorts in the PSID (p. 1002). A slightly higher estimate comes from Ludwig and Miller (2007), who estimate that children's enrollment in Head Start was 17 percentage points higher at the 300-poorest county discontinuity in the 1988 National Educational Longitudinal Survey (NELS).

Based on this evidence, we use our best estimate of 0.149 to transform the ITT effects from our event-study and spline specifications into average treatment-effects-on-the-treated (ATET). We also construct confidence intervals using a parametric bootstrap procedure with 10,000 independent draws from normal distributions with means and standard deviations equal to the point estimates and standard errors from the reduced-form and first-stage estimates (Efron and Tibshirani, 1993).

2.5.2 Did Head Start's Launch Correspond to Other Policy Changes?

Another key assumption of our analysis is that the launch of a Head Start program is the only reason why a cohort's adult outcomes changed for age-eligible children relative to 6+ year olds. The decentralization of U.S. governance and the process of applying for OEO grants makes it unlikely that communities across the nation coordinated in their grant applications. More worrisome is that OEO could have provided multiple grants to a community in the same year, making it difficult to disentangle the effects of Head Start from other OEO programs. To test this hypothesis, we use information on other War on Poverty programs compiled from the National Archives and estimate regressions similar to equation 2.1. In particular, we replace the dependent variable with an indicator for receiving funding in county c in fiscal year t, or $Y_{ct} = \theta_c + \delta_{s(c)t} + X'_{ct}\beta + \sum_k^K \phi_k 1(t - T_c^* = k) + \epsilon_{cbt}$. We also include county and state-year fixed effects, θ_c and $\delta_{s(c)t}$, and county level covariates, X_{ct}, as we do in our primary regression specifications. Our variable of interest is event-time, $t - T_c^* = k$, the year of observation relative to the date Head Start launched (the year before Head Start began, $k = -1$, is omitted).

Figure 2.3 shows the relationship between the launch of Head Start and the launch of other OEO programs. As expected, 100 percent of treated counties in our sample first received a Head Start grant in event-year 0. This is by design. In subsequent years, the share of counties receiving a Head Start grant tapers off to 70 percent after around five years – this reflects the fact that some counties received multi-year grants and also the fact that not all of the early programs continued.

For our estimates to be confounded by changes in other federal funding, these funding changes would need to happen around the time Head Start launched, or event-year 0. However, our analysis of Food Stamps, Community Health Centers, and other child health programs show no such pattern. The one program that shows a small change in funding after Head Start began is the CAP health project. Importantly for our inferences, funding was very small for this program (around \$8 in 2013 dollars per person; by contrast, annual Head Start funding was nearly \$1,500 per 4-year-old during the same period). Furthermore, there is little reason to think this program would affect children who were 5 or younger but not those who were 6. Although we cannot rule out funding changes in programs we do not measure, these patterns provide little evidence that our research design is confounded by other OEO programs.

2.6 Head Start's Effects on Human Capital

Figure 2.4 plots the event-study estimates for all outcomes in the human capital index for the set of compositionally balanced county-birth-cohorts, or individuals ages 15 to as young as -1 when Head Start launched. The solid line plots the event-study estimates. Consistent with the patterns anticipated in Figure 2.2, the human capital index and each of its components exhibit little relationship to adult outcomes for cohorts ages 6 to 15 when Head Start launched (i.e., there is little evidence of a pre-trend among ineligible cohorts). However, the index and many of its subcomponents exhibit a trend-break around age 6, suggesting that access to Head Start improved the human capital of adults. Notably, this relationship is not as sharp as in regression kink designs, because (1) school age-entry cutoffs were not strictly enforced, (2) older children ages 6 and 7 participated in Head Start (although at lower rates), and (3) older siblings of participants may have benefited indirectly from their younger sibling's involvement. Any benefits of Head Start for these other ages would lead the event-study estimates for ages higher than 6 to fall slightly below zero, as the data show. They also weaken the visual evidence and formal tests for a trend-break at exactly age 6.

Table 2.1 summarizes both the event-study and spline estimates at -1. Column 1 presents the mean and standard deviation of the outcome for cohorts ages 6 and 7 at the time of Head Start's launch (our control group). Column 2, which shows the ITT-event-study estimate at -1 (our estimate for cohorts that were age eligible for up to three years for a fully implemented program), suggests that Head Start significantly improved adult human capital. The standardized index increases by 1.5 percent of a standard deviation for the fully exposed cohort and 10 percent of a standard deviation for treated children (column 6). Across outcomes, column 3's ITT-spline estimates are identical to those in column 2 to the hundredth.

Supporting the impression left by the event-study plots, column 4 shows that a formal test for pre-trend (ages 6 to 15), the slope estimate for the spline component for ages 6 to 15). Figure 2.5 additionally plots the magnitude of the t-tests for the pre-trend. For the index and for each of its subcomponents, the data find no evidence of a pre-trend – a conclusion strengthened by Bonferroni-Holm (BH) p-value adjustments for multiple hypothesis testing. T-tests for pre-trends fail to reject the slope is zero in all cases. However, column 5 of Table 2.1 and Figure 2.5 shows that the data reject the null hypothesis of no trend-break at age 6 for the adult human capital index at the 1-percent level. With the exception of high school graduation (which is just below the threshold for statistical significance at the 10-percent level), the data show evidence of a trend break for each of the subcomponents of the human capital index. It appears that , even with Head Start's spill-overs to children older than 6, the relationship between adult human capital changed for children age-eligible for Head Start relative to children in the same county who were old enough to enroll in first grade.

The data show strong effects on some of the most commonly studied outcomes in the preschool literature, including both high school graduation and college enrollment. Figure 2.4B and Table 2.1 show that treated children were 1.9 percentage points more likely to complete high school/GED (column 6) – a 2.1-percent increase relative to the control mean (column 7). The magnitude of this estimate is precisely estimated, but smaller than other estimates of Head Start's effects in the literature. Figure 2.6A shows that the effect is roughly half the size of Garces, Thomas and Currie (2002)'s sibling comparison in the PSID and Thompson (2017)'s spending design in the NLSY. In addition, it is one-fifth the size of Johnson and Jackson (2017)'s spending design estimates for the very disadvantaged sample in the PSID; and one-ninth the size of Ludwig and Miller (2007)'s RD estimates using the Census. (It is one quarter the size of Deming (2009)'s sibling comparison for Head Start in the 1990s for more recent cohorts.) Although our estimate falls within the confidence intervals of previous studies, this reflects the imprecision of those estimates.

Figure 2.4C and Table 2.1 also shows a statistically significant effect of Head Start on college enrollment. Head Start raised college enrollment by 5.4 percentage points, or 8.7 percent. This estimate is half the size of Garces, Thomas and Currie (2002) and one quarter the size of Ludwig and Miller (2007) (Figure 2.6B). (The magnitude of the increase in college enrollment of 0.05 is only slightly smaller than Deming (2009)'s NLSY sibling comparison for Head Start in the 1990s.) Again, consistent with the visual impression of a trend break in Figure 2.4C, we find no evidence of a pre-trend for children older than 6 at launch and reject the null hypothesis of no trend break at age 6 at the 5-percent level.

In addition to generating more precise estimates for these commonly studied outcomes, our large-scale data permit a novel evaluation of the effects of Head Start on other dimensions of human capital, including college completion or higher degrees, which previous data have not been able to detect. Table 2.1 shows that participating children achieved 19 percent higher college graduation rates (the trend break is statistically significant at the 1-percent level). These estimates are one-quarter to one-fifth the size of those found for the Abecedarian Project (Currie, 2001; Barnett and Masse, 2007; Duncan and Magnuson, 2013). Similarly, completion of professional or doctoral degrees increased by 50 percent among treated children, although evidence of a significant pre-trend caution against a causal interpretation (this is also consistent with Figure 2.4F). These gains across the education distribution are summarized in a 0.29-year increase in schooling. This estimate is smaller than Johnson and Jackson (2017)'s estimate of 0.52 years for very disadvantaged children, but it is highly statistically significant and is not driven by a pre-trend.

These large effects on college and higher degrees may be surprising, given that no other study of preschool has documented effects on post-secondary education. This lack of evidence may reflect, in part, the small longitudinal samples or the small scale of model preschool programs. Differences in the participating children may also matter. Abecedarian and Perry's participants were very disadvantaged children and mostly black, and Perry's participants had low IQs. In contrast, Head Start was not exclusively for poor, African-American, or low-IQ children. Consequently, Head Start's participants in the 1960s and 1970s likely faced fewer socio-economic and cognitive disadvantages and less racism relative to model programs. Differences in the background characteristics of Head Start's participants make it less surprising that they experienced gains in post-secondary education.

Because analyses of model preschool programs have found different educational effects for boys, Table 2.2 stratifies our sample by sex. Among participating men, the human capital index increased by a statistically significant 14 percent of a standard deviation. For this group, high school completion rose by a statistically insignificant 2.7 percent, college attendance rose by 13 percent, and college completion rose by 27 percent. The high school estimates are smaller than others in the literature, but the college attendance estimates tend to be larger. Head Start cumulatively raised years of education among treated men by 0.41 years and the likelihood of completing a professional/doctoral degree by 59 percent. The evidence suggests that men treated with Head Start were 19 percent more likely to hold professional jobs.

The human capital index increased by less among women, at only 7 percent of a standard deviation. Completion of high school (or a GED) rose by a statistically insignificant 1.5 percent, and college attendance rose by 5.7 percent (although the trend-break is not statistically significant). For women, changes in the human capital index appear driven by increases in higher degrees, including an 11-percent increase college completion and 36-percent increase in professional degrees. Treated women's schooling rose by 0.17 years and their likelihood of holding a professional job rose by 9.5 percent.

Our Appendix Tables B.7-B.8 report estimates of Head Start's effects on human capital by race. Unfortunately, even our large sample size is too small to precisely quantify effects by race, because less than 1/6 of our sample is nonwhite and, unlike longitudinal data, we have no background covariates to model the many other determinants of adult outcomes. The broad patterns in these estimates suggest that Head Start's effects are largest among white men (13 percent of a standard deviation) and smaller among white and non-white women (5-6 percent of a standard deviation, respectively). Effects for non-white men were generally small and imprecise.

2.7 Head Start's Effects on Economic Self-Sufficiency

The substantial effects of Head Start on human capital suggest a potential for effects on economic self-sufficiency. Figure 2.7A plots the event study estimates for the self-sufficiency index, and Table 2.3 shows that an index of economic self-sufficiency aggregated over both sexes was by 4 percent of a standard deviation higher than for children ages 6-7 at the time Head Start began. Consistent with Head Start affecting less skilled individuals, the program decreased the likelihood of adult poverty by 12 percent and receipt of public assistance income by 29 percent, though these results are imprecise. Figure 2.7B and 2.7C show striking evidence of a trend break at age 6, which is also reflected in column 5 of Table 2.3 and in Figure 2.6B. However, there is little effect of Head Start on labor-force participation or wage income. (We omit the event study estimates for these outcomes for parsimony, because they are noisy and show no effect.) Null effects for wages and labor-force participation also affect the magnitude of the change and the trend-break in the economic self-sufficiency index, which does not exhibit the sharp trend-break at age 6 as in the human capital outcomes.

This result may reflect the fact that men's and women's work effort changed in offsetting ways, resulting in selection for both groups. Whereas Head Start's effect on men's human capital may have led them to increase employment (e.g., the substitution effect dominates), the reverse may be true for women (e.g., the income effect dominates as more education allows them to marry higher-earning men).

Table 2.4 stratifies our sample by sex and provides evidence consistent with this hypothesis. Because the self-sufficiency estimates are noisier and stratifying by sex reduces sample sizes, we focus our discussion on the spline estimates. For treated men, the self-sufficiency index increased by 3 percent of a standard deviation. We also find positive effects of Head Start exposure on both the extensive and intensive margins of men's labor-force participation. Treated men were 2.1 percent more likely to have worked for pay (column 7), worked an average of one more week and one more hour per week (column 6). Consistent with these estimates reflecting the causal effect of Head Start, we find no evidence of a pre-trend and a marginally significant trend-break at age 6, which does not survive the Bonferroni-Holm standard error correction. At first glance, it is curious that the combined effects of increased human capital and labor-force participation do not appear to have affected annual wages. Upon further investigation, this appears consistent with Head Start inducing negative selection into the labor-force: the marginal participants tended to be less skilled, and therefore lowered the cohort's wages on average. Head Start had little effect on men's poverty, but the program is associated with a 27-percent decline in public assistance receipt among treated men. Reductions in public assistance are also consistent with negative selection, because male public assistance recipients receive high rates of disability income.

The pattern is different for women. The self-sufficiency index increased by 4 percent of a standard deviation among women treated with Head Start, largely driven by a 28-percent reduction in public assistance receipt and a 16-percent reduction in poverty. However, women's labor-force participation on the extensive and intensive margins fell slightly, albeit not significantly. These reductions in work appear to have increased annual wages of working women by around 4 percent, which is consistent with Head Start inducing positive selection (e.g., less-skilled women opting out). In our conclusion, we examine the effects of negative selection among men and positive selection among women.

As with the human capital outcomes, our Appendix Tables B.10-B.11 report estimates of Head Start's effects on self-sufficiency by race and sex. Similarly, the effects for nonwhites are generally statistically insignificant, owing to the fact that nonwhites comprise less than 15 percent of the sample and that we have few background characteristics to model the considerable variation in outcomes. As with human capital, the patterns of these estimates suggest that Head Start's effects on self-sufficiency are largest among white men and women (3 and 4 percent of a standard deviation, respectively).

2.8 Heterogeneity in Head Start's Long-Run Effect

This final section seeks to shed light on the potential mechanisms for Head Start's effects by examining how the estimates vary with access to other public programs and local economic conditions. We implement this analysis by interacting a binary indicator for whether cohorts lived in counties with "high" or "low" exposure to a program with the spline components in equation (2), where "high" is equal to one for counties above the median in the characteristics and 0 otherwise. We use only the main indices for human capital and economic self-sufficiency as the dependent variables. We caution that the lack of randomization of alternative programs and conditions means that, while these relationships are suggestive, they should not be interpreted as causal. Additionally, uncertainty about how program enrollment varied means these estimates are less precise than desired. We nevertheless provide additional guidance about the magnitudes of the ATET effects using our estimates of differential take-up across locations to scale our ITT estimates.

We first investigate the hypothesis that Head Start's long-run effects relate to their complementarities with other health programs for disadvantaged children. If health screening and referrals to health services (a sizable share of Head Start's budget) played a role in driving long-term effects, we would expect Head Start's effects to be larger for children with greater access to these health services through community health centers (CHCs) and/or Medicaid. (We would also expect the selection effects on wages to mirror those we observe for men in Table 2.4.) Table 2.5 provides suggestive evidence of this mechanism, showing that the ATETs of Head Start for human capital were slightly larger in areas with CHCs and three times as large in states where more children were eligible for Medicaid (17 percent increase relative to 5 percent for less-exposed children). The 95-percent confidence interval in the difference between these two treatment effects on the treated suggests we can reject the null hypothesis of equal effects. Consistent with health services bringing more previously disabled workers into the labor market, Head Start's ATETs on economic self-sufficiency are more muted in locations with greater health services.

A related hypothesis is that Head Start affected adult outcomes by providing healthy meals and snacks, improving child nutrition which increased both health and learning. If nutrition is an important mechanism for Head Start's long-run effects, we would expect the program's effects to be smaller for children with greater access to Food Stamps, which also supported the provision of healthy meals. Consistent with this, Table 2.5 shows that children participating in Head Start with more access to Food Stamps experienced smaller – although still statistically significant – gains in human capital and economic self-sufficiency. This suggests that the Food Stamps program provided a partial substitute for Head Start's nutritional component.

The OEO's larger effort to set up Head Start programs in the poorest 300 counties could also lead the Head Start program to be more intensive in these areas, potentially having larger effects. For this test, we report the poorest 300 counties in the column for "above median" and report the effects for counties outside this group in the column "below median." While Table 2.5 shows little evidence of differential effects on human capital across these two groups of counties, there is suggestive evidence that children in poorer counties benefitted more in terms of their economic self-sufficiency-although this difference is imprecise.

A final hypothesis is that Head Start's effects should be larger in areas with greater subsequent economic growth. Strong economic conditions should both increase the resources of children's parents, the provision of public goods (such as schools), and create stronger incentives for children to invest in themselves, as children could expect higher and more certain returns. Rather than using actual economic growth (that may be endogenous), we use predicted economic growth between 1965 and 1985. Table 2.5 suggests that the benefits of a strong economy complement Head Start's effects. The ATETs of Head Start for human capital were twice as large in areas with strong predicted economic growth than in areas with weaker predicted economic growth. The ATETs of Head Start for economic self-sufficiency were fifty percent larger large in areas with strong predicted economic growth.

All in all, these results suggest that Head Start's long-run effects may be driven by many factors beyond a preschool curriculum, including health screenings and referrals and more nutritious meals for an a population thatwith otherwise may have been under-nourished and had little access to health care and under-nourished. Unsurprisingly, the effects of Head Start appear to be complementary to the family and public resources arising from a stronger economy.

2.9 New Evidence on the Long-Term Returns to Head Start

Over the past 20 years, substantial evidence has accumulated that model preschool programs have sizable economic returns (Almond and Currie, 2011; Duncan and Magnuson, 2013; Heckman et al., 2010; Cunha and Heckman, 2007). However, convincing evidence on the long-run returns to larger-scale, public preschool has remained sparse (Phillips et al., 2017).

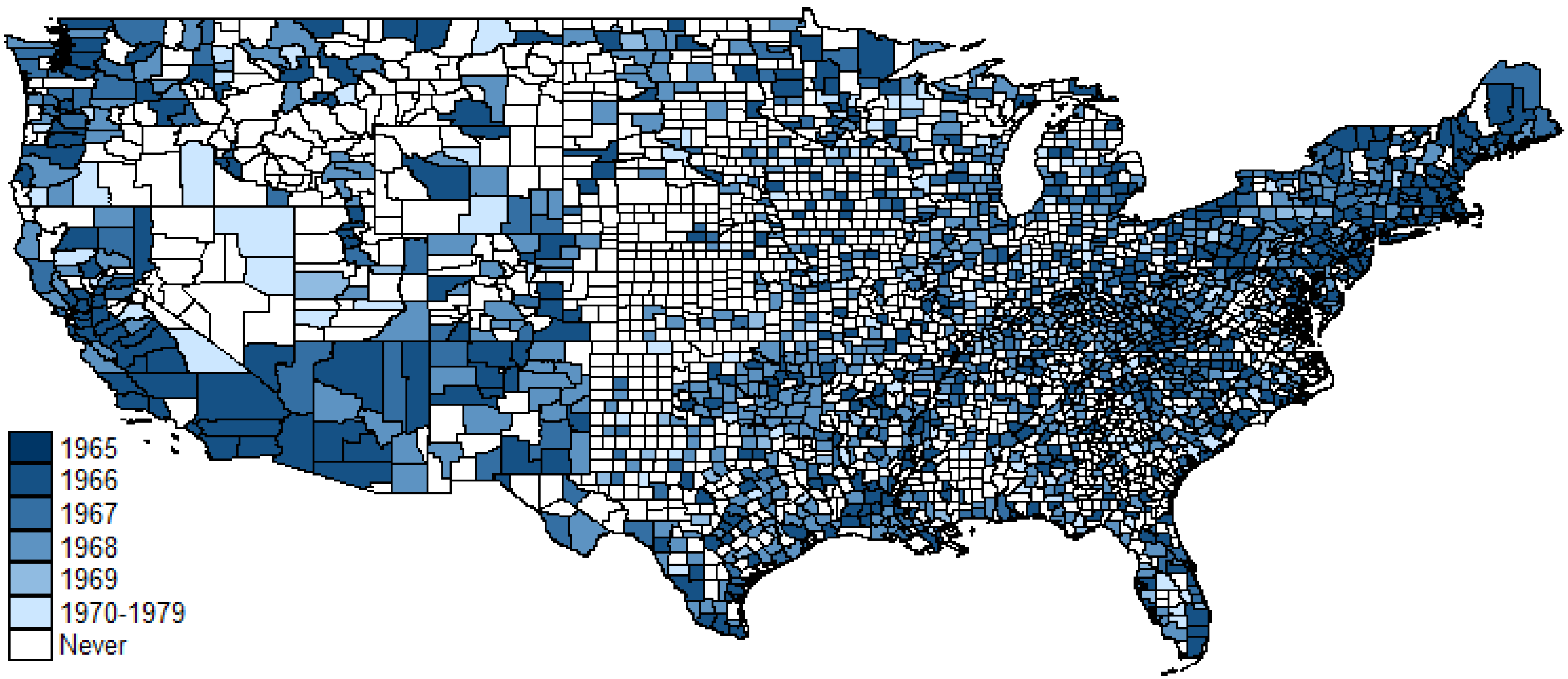
Using large-scale restricted Census/ACS data, this paper provides new evidence of the long-term effects of Head Start, the nation's longest-running, large-scale public preschool program. We find that Head Start had large effects on participants' human capital. Head Start children achieved 0.29 more years of schooling, reflecting the fact that they were 2.1 percent more likely to complete high school, 8.7 percent more likely to enroll in college, and 19 percent more likely to complete college. A second finding is that Head Start increased adult self-sufficiency, reducing the likelihood of adult poverty by 12 percent and public assistance receipt by 29 percent. Heterogeneity tests suggest that these long-run effects may reflect many aspects of the Head Start program beyond its curriculum: health screenings and referrals and more nutritious meals appear to be important mechanisms for the program's effects on disadvantaged children. In addition, the effects of Head Start appear to be complement greater family and public resources arising from a stronger economy.

A full accounting of the costs and benefits of Head Start is beyond the scope of this paper, but we summarize the implications of our estimates using potential earnings to account for selection (Neal and Johnson, 1996; Deming, 2009). Following Neal and Johnson (1996) and Deming (2009), the advantage of this approach is that it allows us to account for the effects of Head Start on employment (which differed for women and men). Like Deming, we use the NLSY79 to predict wages for individuals born from 1957 to 1965 (ages 14 to 22 in 1979) – a time frame that overlaps our Census/ACS analysis. The NLSY data allow us to estimate the relationship between wages and components of the human capital and economic self-sufficiency indices after flexibly controlling for ability using the AFQT. Although AFQT is not available in the Census/ACS, using this as a covariate helps mitigate omitted variables bias in ability in the education and earnings relationship. These regressions are reported in our Appendix Table B.6. After accounting for ability, the NLSY79 suggest a private internal rate of return to Head Start of 7.7 percent, which ranges from 4 percent for women to 11 percent for men. As an alternative, the internal rate of return of putting one child through Head Start is 2.4 percent using only savings on public assistance expenditures (estimated at $9,967 in the Survey of Income and Program Participation).

Several reasons suggest that these estimates are conservative. First, our research design differences out sibling spill-over effects, which tends to reduce the estimated effect sizes. Second, reports of income and public assistance receipt may be severely underreported in major national surveys (Meyer, Mok and Sullivan, 2015; Bound, Brown and Mathiowetz, 2001), suggesting estimates of Head Start's effect on public assistance may be understated. Third, adding increases in tax revenues and, reductions in deadweight loss from public assistance transfers, or underreporting in public assistance income would serve to increase our estimates of the returns to Head Start. Finally, estimates of the returns to Head Start ignore benefits through improvements in outcomes not measured here. For instance, they ignore the extent to which more education engenders better health, longevity, or well-being. These potential limitations, however, tend to strengthen the conclusion that Head Start achieved its goal of reducing adult poverty, delivering sizable returns to investments made in the 1960s and 1970s. The results suggest potentially larger social returns.

The long-run returns to today's public preschool programs may be different for a number of reasons. Today, the curriculum is different, the target population is different, and the alternative programs and resources available to poor children are radically different than in the 1960s. Of course, researchers will need to wait another 50 years to evaluate the long-run effects of today's preschool programs. In the meantime, the sizable returns to the "less-than-model" Head Start preschool program of the 1960s suggest productive avenues for improving the lot of poor children today.

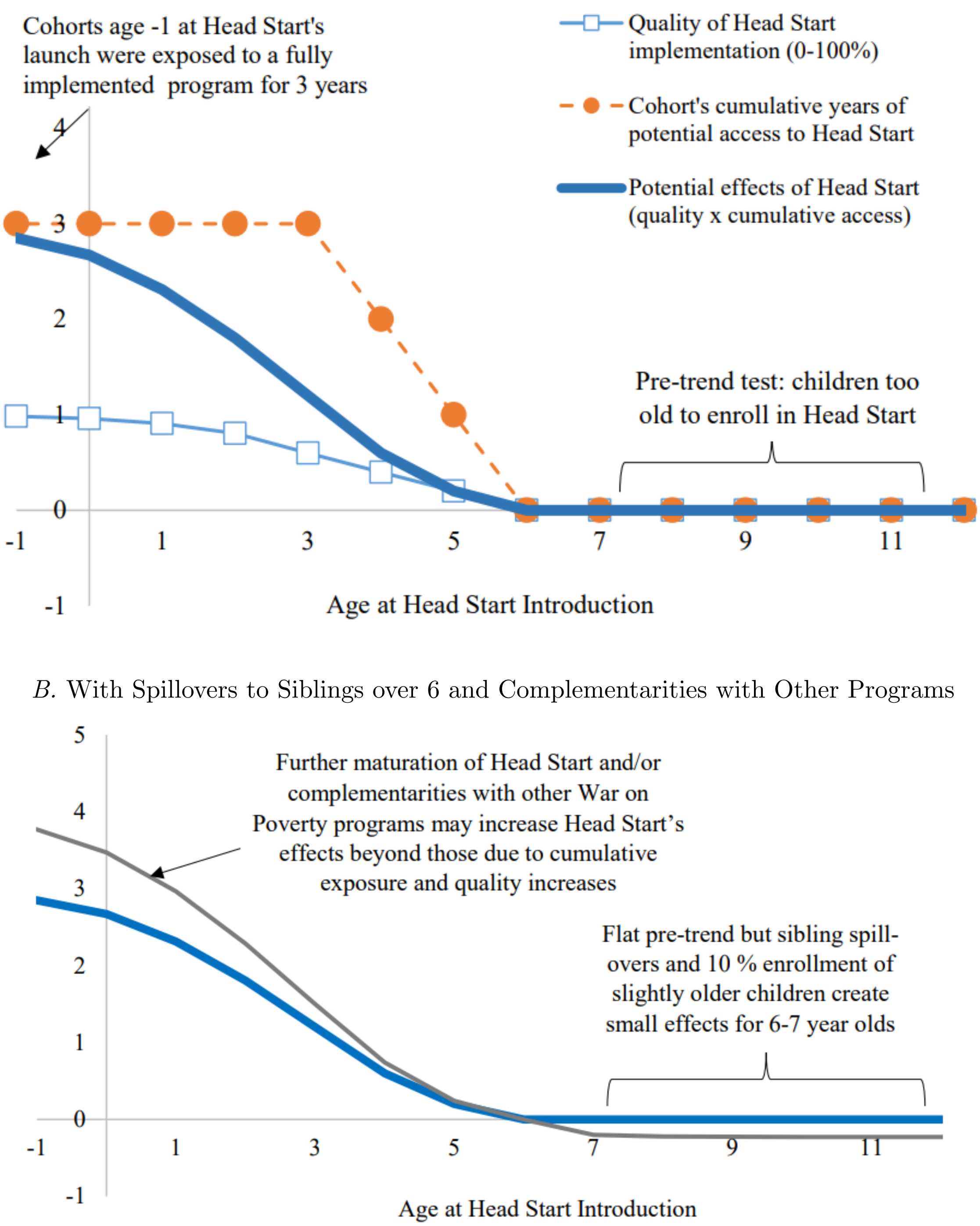
Figure 2.1. The Launch of Head Start Between 1965 and 1980



Notes: Counties are grouped by the fiscal year that Head Start launched between 1965 and 1980. Data on federal grants are drawn from the National Archives and Records Administration (NARA). See Bailey and Duquette (2014) and Bailey and Goodman-Bacon (2015) for details on data and variable construction.

Figure 2.2. The Expected Pattern of Effects on Adult Outcomes by Age of Child at Head Start's Launch

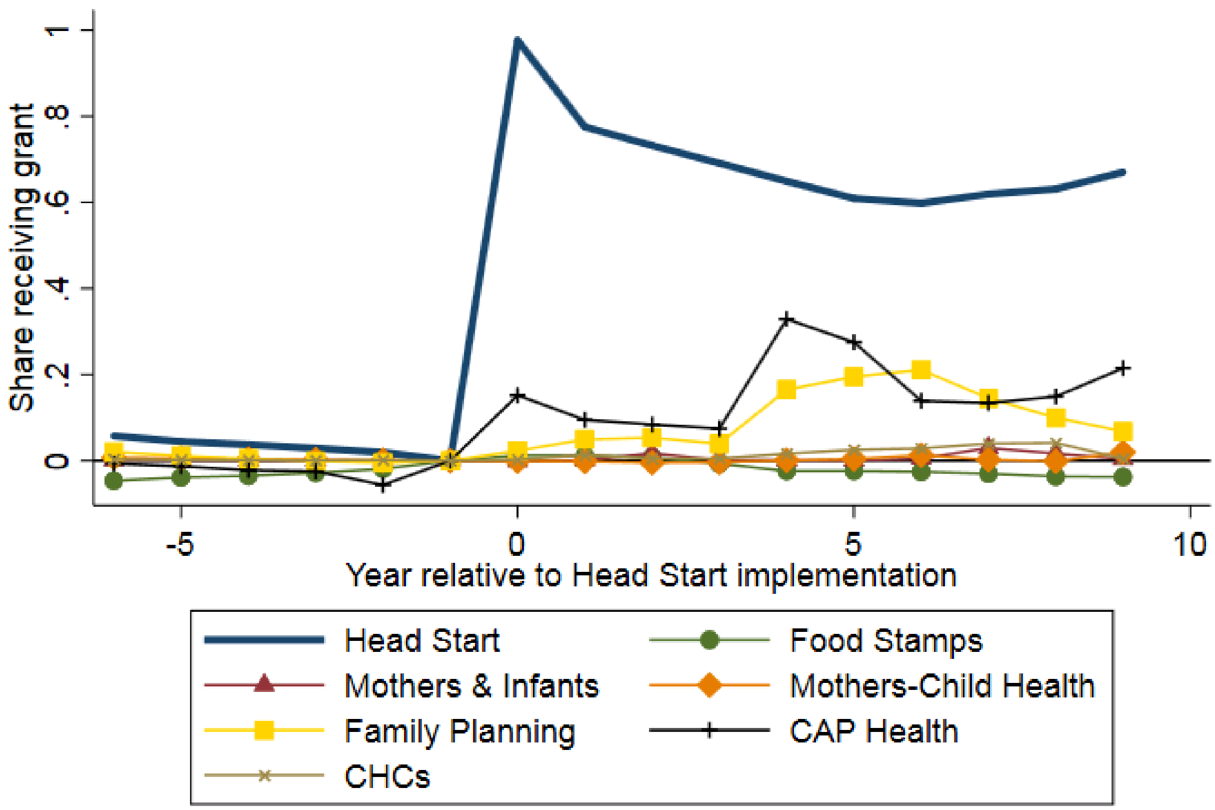
A. No Sibling Spillovers or Complementarities with Other Programs



B. With Spillovers to Siblings over 6 and Complementarities with Other Programs

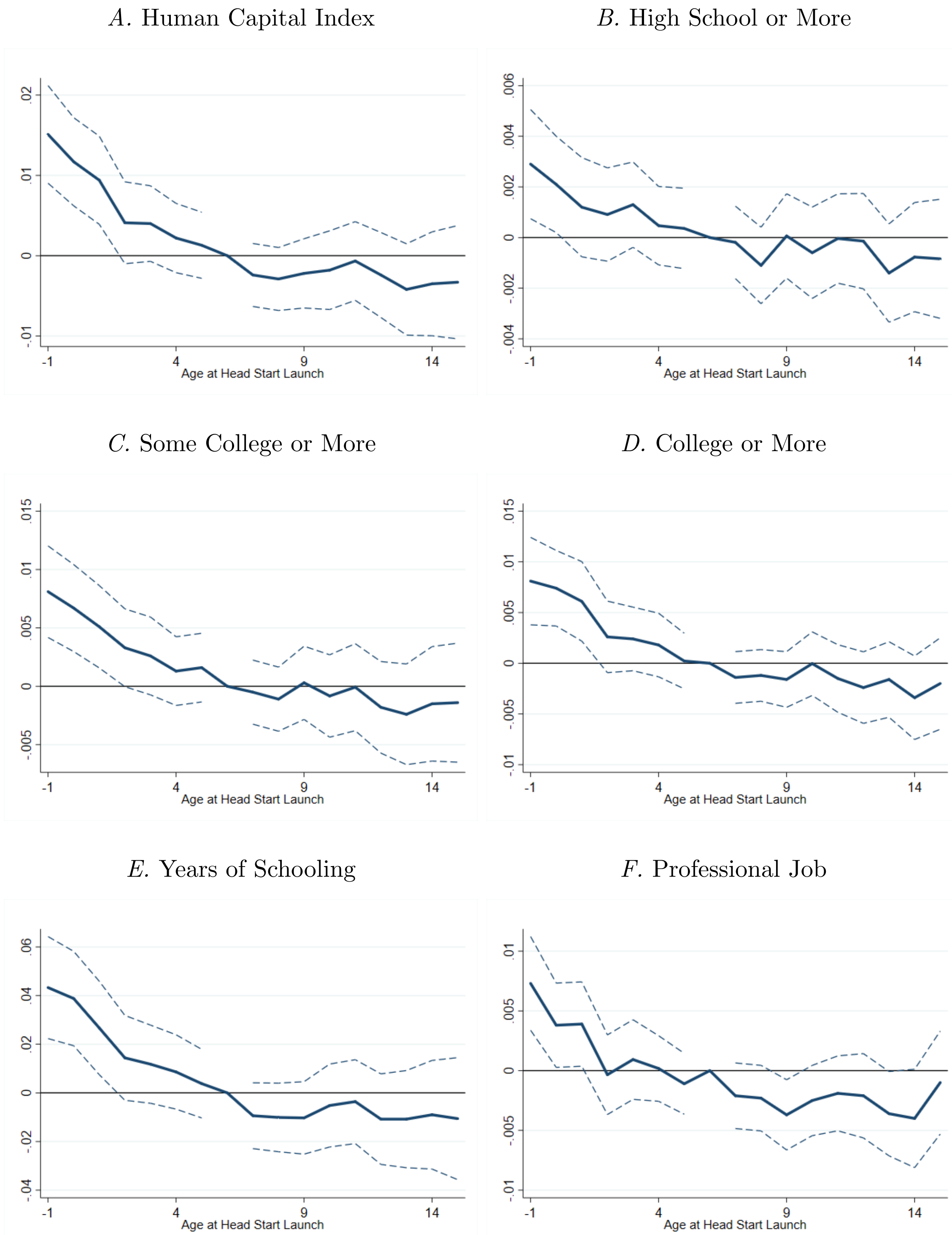
Notes: Figure 2.2A illustrates the potential effects of Head Start assuming there are no effects on children 6 and over, no spillovers to older siblings, and no complementarities with other programs. Figure 2.2B illustrates the way spillovers and complementarities could alter the pattern of the program's effects.

Figure 2.3. Funding for Other OEO Programs Relative to the Year Head Start Began



Notes: Dependent variable are binary variables for whether a county received a grant for the indicated program in the indicated year. Data on federal grants and programs are drawn from the NARA.

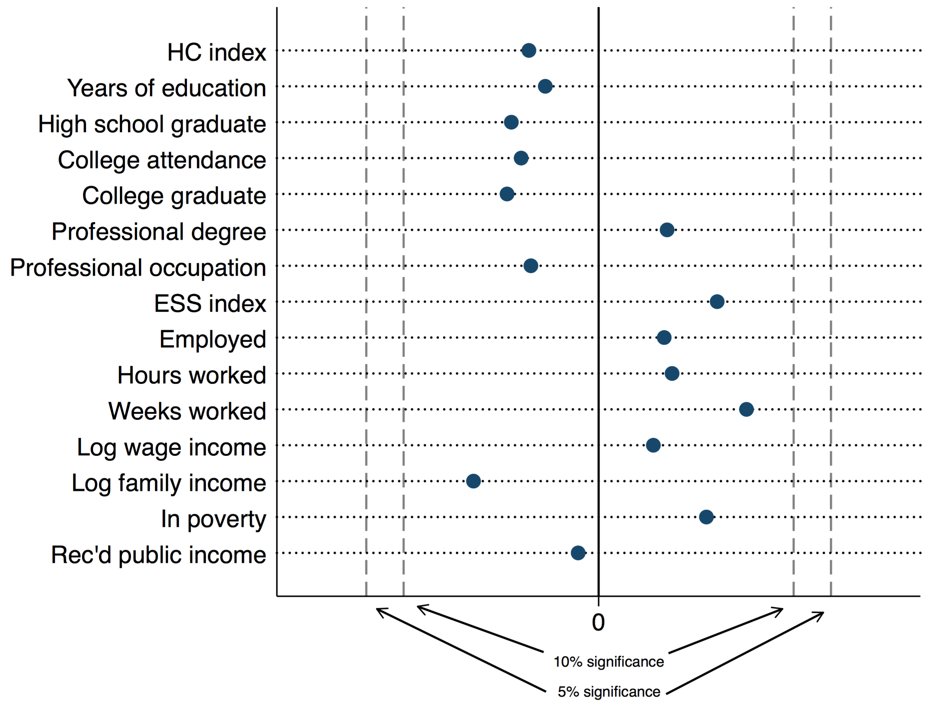
Figure 2.4. The Effect of Head Start on Adult Human Capital



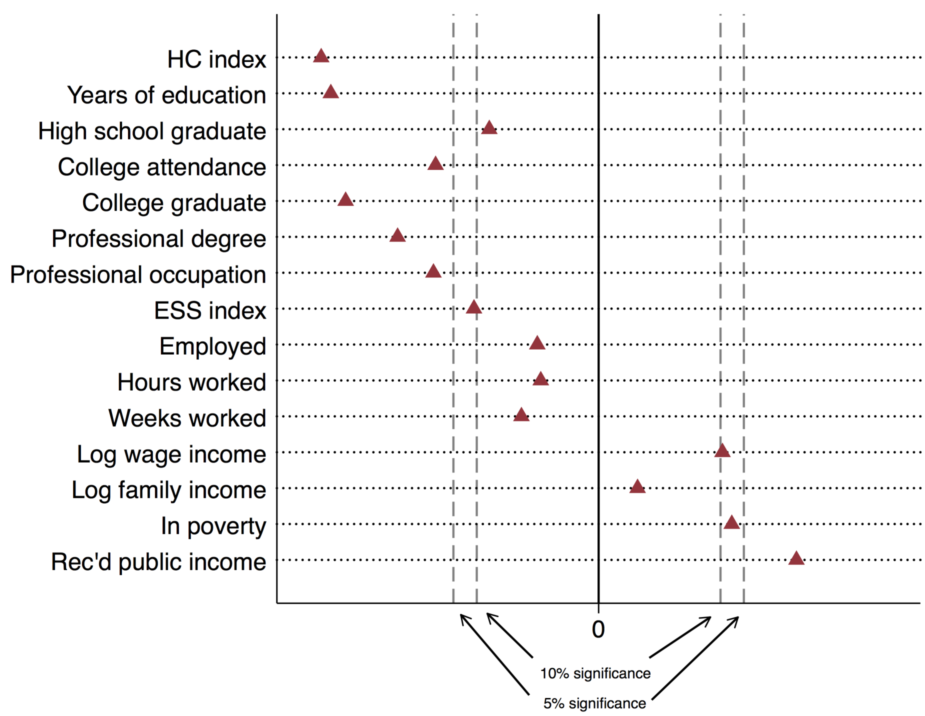
Notes: The figures plot event-study estimates of ϕ for different outcomes using the specification in equation 2.1. Standard errors clustered at the county level. Dashed lines show 95-percent, point-wise confidence intervals for each estimate.

Figure 2.5. Visual Representation of Test Statistics Evaluation Pre-Trends and Trend Breaks

A. Test for Pre-Trend (slope of spline for age 6-15 at Head Start's launch)



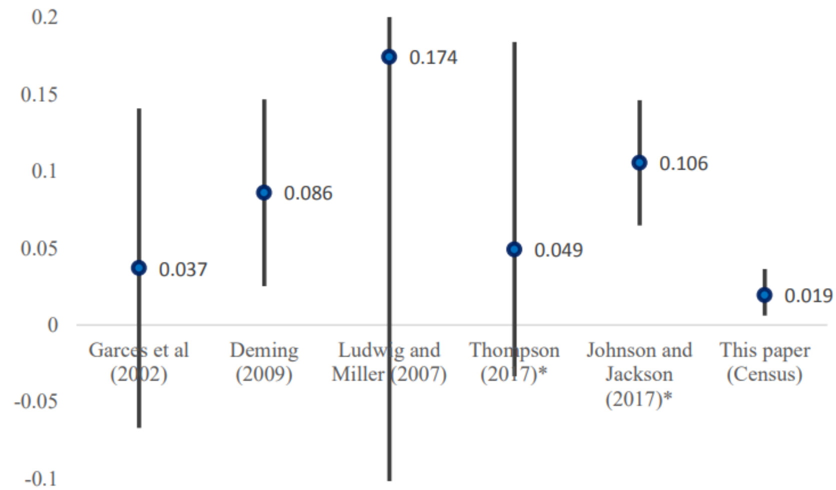
B. Test for Trend Break (change in spline slope at age 6, before which children are age-eligible for Head Start)



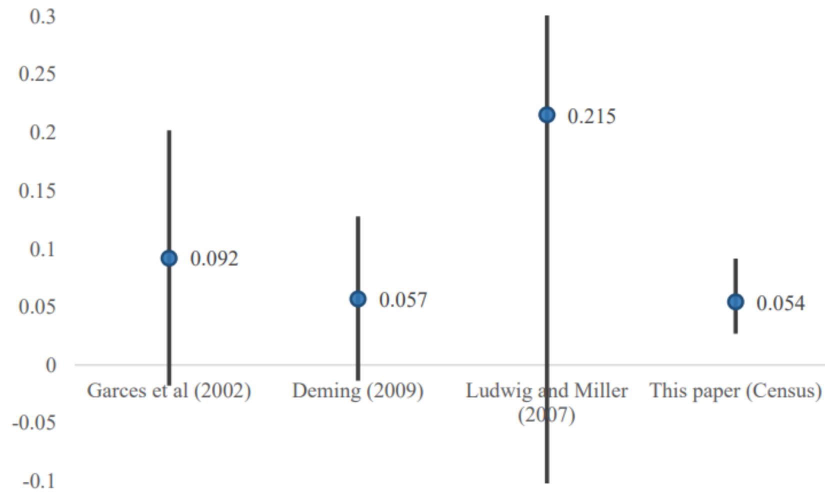
Notes: The figure plots the t-statistic on the slope of the spline for ages 6-15 (panel A) or the F-statistic for the test for a trendbreak at age 6 (panel B). Dashed lines show the threshold for statistical significance at the 10 and 5 percent levels. Compare these to columns 4 and 5 of Tables 2.1 and 2.3.

Figure 2.6. The Magnitude of Head Start's Effects on Education Across Studies

A. Effects of Head Start on High School Graduation



B. The Effects of Head Start on College Enrollment

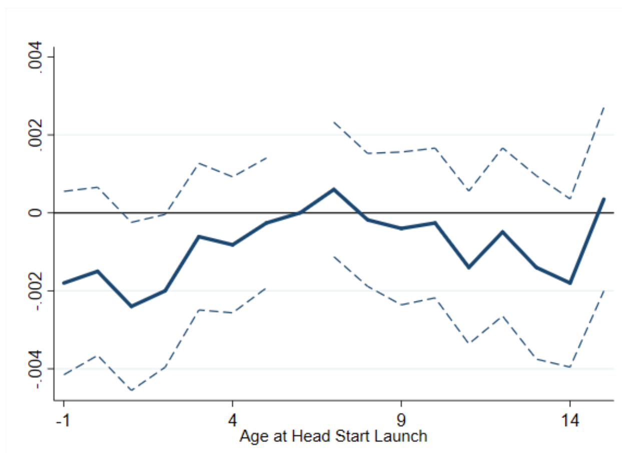


Notes: Circles indicate the reported or derived ATET from different studies. For sibling fixed effect studies, the ATET is directly reported in the papers. Because we cannot resample from data used in other Head Start papers, we calculate the ATETs for other papers using a parametric bootstrap procedure using 10,000 independent draws from normal distributions with means and standard deviations equal to the point estimates and standard errors from the reduced-form and first-stage estimates (Efron and Tibshirani, 1993). Because Johnson and Jackson (2017) does not report a standard error on the first stage, the confidence interval reported for this study in Panel A does not include this first-stage uncertainty. We limited the y-axis range so that the confidence intervals for most studies could be read from the figure. The confidence intervals for Ludwig and Miller (2007) fall outside the y-axis range and are [-0.54,1.47] in panel A and [-0.67,1.82] in panel B. Bars indicate the reported 95-percent confidence interval for sibling fixed-effect models or constructed for the ITT studies as described in the text. See Appendix for more details on the exact figures used. *Johnson and Jackson (2017) and Thompson (2017) sample likely eligible samples of the PSID and NLSY79: individuals born to parents in the bottom quartile of the income distribution, and parents with no college education, respectively.

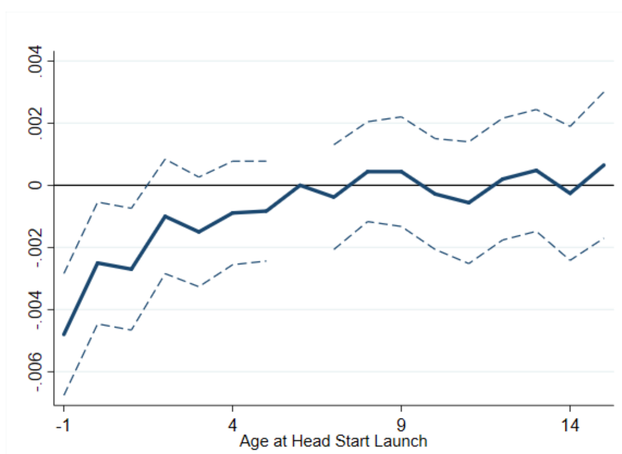
Figure 2.7. The Effect of Head Start on Adult Economic Self-Sufficiency

A. Economic Self-Sufficiency Index

B. In Poverty



C. In Poverty



Notes: See Figure 2.4 notes. Note that In Poverty and Received Public Assistance are reverse coded when included in the Economic Self-Sufficiency Index.

Table 2.1. The Effect of Head Start on Adult Human Capital

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 (p-val) [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
Human capital index	0.014	0.015	0.015	-0.0002	14.02	0.101	
	(0.198)	(0.003)	(0.003)	(0.0003)	(0.00)	[.057,.163]	
Completed high school/GED	0.92	0.003	0.003	-0.0001	2.18	0.0189	2.1%
	(0.078)	(0.001)	(0.001)	(0.0001)	(0.14)	[.005,.038]	
		[0.01]	[0.01]	[1.00]	[0.14]		
Attended some college	0.62	0.008	0.008	-0.0002	4.84	0.054	8.7%
	(0.140)	(0.002)	(0.002)	(0.0003)	(0.03)	[.027,.092]	
		[0.00]	[0.00]	[1.00]	[0.08]		
Completed 4 year college	0.29	0.008	0.009	-0.0002	11.66	0.054	18.6%
	(0.122)	(0.002)	(0.002)	(0.0002)	(0.00)	[.025,.094]	
		[0.00]	[0.00]	[1.00]	[0.00]		
Prof. or doc. degree	0.028	0.002	0.002	0.0000	7.36	0.014	50.0%
	(0.037)	(0.001)	(0.000)	(0.0001)	(0.01)	[.006,.024]	
		[0.00]	[0.00]	[1.00]	[0.03]		
Years of schooling	13.57	0.043	0.049	-0.0005	13.06	0.291	2.1%
	(0.695)	(0.011)	(0.010)	(0.0012)	(0.00)	[.144,.49]	
		[0.00]	[0.00]	[1.00]	[0.00]		
Has a professional job	0.35	0.007	0.007	-0.0001	4.97	0.0489	14.0%
	(0.121)	(0.002)	(0.002)	(0.0002)	(0.03)	[.022,.085]	
		[0.00]	[0.00]	[1.00]	[0.08]		

Notes: In column 1, the control mean and standard deviation are calculated using the cohorts ages 6 and 7 at the time Head Start was launched. Column 2 presents the estimated intention-to-treat (ITT) effect evaluated at birth cohort of full exposure (-1, see Figure 2.2). Column 3 presents the ITT spline estimate evaluated at -1. Column 4 presents the pre-trend estimate for the spline segment for age 6 and older at implementation. Column 5 presents the F-statistic and p-value for the test of a trend-break in the spline at age 6. The ATET estimate in column 6 divides the ITT effect at -1 by the estimate of receiving a Head Start grant on school enrollment at school age 5, 0.149 (s.e. 0.022) for the full sample and 0.151 (s.e. 0.022) for men and 0.145 (s.e. 0.022) for women; see Appendix Table B.5). Column 7 computes the percentage increase implied by the ATET relative to the control mean (the ratio of Column 6 to Column 1) for components of the index. The BH p-values presented in columns 2, 4, and 5 in brackets use the Bonferroni-Holm method to account for multiple hypotheses testing of individual outcomes within an index.

Table 2.2. The Effect of Head Start on Adult Human Capital by Sex

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 (p-value) [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
A. Men							
Human capital index	0.014	0.021	0.02	-0.0005	11.79	0.1360	
	(0.247)	(0.004)	(0.003)	(0.0004)	(0.00)	[.081,.215]	
Completed high school/GED	0.91	0.004	0.003	-0.0001	1.92	0.0250	2.7%
	(0.104)	(0.002)	(0.001)	(0.0002)	(0.17)	[.005,.049]	
		[0.01]	[0.01]	[1.00]	[0.22]		
Attended some college	0.59	0.011	0.01	0.0000	5.67	0.0740	12.5%
	(0.176)	(0.003)	(0.002)	(0.0003)	(0.02)	[.036,.125]	
		[0.00]	[0.00]	[1.00]	[0.07]		
Completed 4 year college	0.29	0.012	0.011	-0.0003	9.99	0.0769	26.5%
	(0.150)	(0.003)	(0.002)	(0.0003)	(0.00)	[.041,.126]	
		[0.00]	[0.00]	[1.00]	[0.01]		
Prof. or doc. degree	0.034	0.003	0.003	0.0000	4.73	0.0199	58.5%
	(0.052)	(0.001)	(0.001)	(0.0001)	(0.03)	[.008,.035]	
		[0.00]	[0.01]	[1.00]	[0.09]		
Years of schooling	13.5	0.062	0.063	-0.0012	12.15	0.4129	3.1%
	(0.878)	(0.013)	(0.011)	(0.0015)	(0.00)	[.228,.668]	
		[0.00]	[0.00]	[1.00]	[0.00]		
Has a professional job	0.34	0.01	0.009	-0.0006	2.54	0.0649	19.1%
	(0.153)	(0.003)	(0.002)	(0.0003)	(0.11)	[.031,.11]	
		[0.00]	[0.00]	[0.06]	[0.22]		
B. Women							
Human capital index	0.015	0.01	0.012	0.0000	6.50	0.0659	
	(0.228)	(0.004)	(0.003)	(0.0004)	(0.01)	[.014,.132]	
Completed high school/GED	0.93	0.002	0.002	-0.0001	1.42	0.0140	1.5%
	(0.087)	(0.001)	(0.001)	(0.0001)	(0.23)	[-.003,.035]	
		[0.28]	[0.07]	[1.00]	[0.47]		
Attended some college	0.65	0.005	0.006	-0.0004	1.20	0.0370	5.7%
	(0.163)	(0.002)	(0.002)	(0.0003)	(0.27)	[.005,.077]	
		[0.15]	[0.01]	[0.97]	[0.47]		
Completed 4 year college	0.29	0.005	0.008	-0.0002	4.63	0.0320	11.0%
	(0.147)	(0.003)	(0.002)	(0.0003)	(0.03)	[-.004,.076]	
		[0.28]	[0.01]	[1.00]	[0.16]		
Prof. or doc. degree	0.022	0.001	0.001	0.0001	3.78	0.0080	36.4%
	(0.042)	(0.001)	(0.001)	(0.0001)	(0.05)	[-.001,.018]	
		[0.28]	[0.07]	[0.91]	[0.16]		
Years of schooling	13.6	0.024	0.038	-0.0003	4.91	0.1659	1.2%
	(0.795)	(0.013)	(0.012)	(0.0014)	(0.03)	[-.013,.379]	
		[0.28]	[0.01]	[1.00]	[0.16]		
Has a professional job	0.37	0.005	0.005	0.0004	4.35	0.0350	9.5%
	(0.150)	(0.003)	(0.002)	(0.0003)	(0.04)	[0,.078]	
		[0.25]	[0.07]	[0.97]	[0.16]		

Notes: See Table 2.1 notes.

Table 2.3. The Effect of Head Start on Adult Economic Self-Sufficiency

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 [p-val] [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
Self-sufficiency index	0.024	0.0055	0.0048	0.0003	4.24	0.037	
	(0.154)	(0.0024)	(0.0022)	(0.0002)	(0.04)	[.005,.076]	
Worked last year	0.86	0.00082	0.00091	0.0000	0.22	0.006	0.7%
	(0.098)	(0.0014)	(0.0013)	(0.0002)	(0.64)	[-.013,.026]	
		[1.00]	[1.00]	[1.00]	[1.00]		
Weeks worked last year	41.1	0.047	0.055	0.0057	0.61	0.32	0.8%
	(5.281)	(0.075)	(0.0721)	(0.0092)	(0.43)	[-.68,1.4]	
		[1.00]	[1.00]	[1.00]	[1.00]		
Usual hours works per week	35.72	0.0334	0.056	0.0106	1.09	0.22	0.6%
	(4.907)	(0.080)	(0.0791)	(0.0085)	(0.30)	[-.84,1.4]	
		[1.00]	[1.00]	[1.00]	[1.00]		
Log labor income	10.6	0.0055	0.0061	0.00018	2.8	0.037	
	(0.268)	(0.0035)	(0.0031)	(0.0004)	(0.09)	[-.009,.09]	
		[0.70]	[0.30]	[1.00]	[0.47]		
Log family income/poverty	5.86	0.0059	0.0054	-0.0004	0.28	0.04	
	(0.258)	(0.0043)	(0.0041)	(0.0004)	(0.60)	[-.017,.104]	
		[0.70]	[0.75]	[1.00]	[1.00]		
In poverty*	0.1	-0.0018	-0.002	0.0001	3.23	-0.012	-12.0%
	(0.084)	(0.0012)	(0.0011)	(0.0001)	(0.07)	[-.03,0.004]	
		[0.70]	[0.35]	[1.00]	[0.44]		
Rec'd public assistance*	0.11	-0.0048	-0.0037	0.0000	7.13	-0.032	-29.1%
	(0.093)	(0.0010)	(0.0008)	(0.0001)	(0.01)	[-.052,-.018]	
		[0.00]	[0.00]	[1.00]	[0.05]		

Notes: *In poverty and received public program income are reverse-coded when used in the self-sufficiency index. See also Table 2.1 notes.

Table 2.4. The Effect of Head Start on Adult Economic Self-Sufficiency by Sex

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 [p-val] [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
A. Men							
Self-sufficiency index	0.031	0.0046	0.00553	-0.0001	4.24	0.0299	
	(0.201)	(0.0031)	(0.0025)	(0.0003)	(0.21)	[-.01,.077]	
Worked last year	0.91	0.0029	0.00301	0.0000	2.33	0.0189	2.1%
	(0.108)	(0.0014)	(0.0011)	(0.0001)	(0.13)	[.001,.041]	
		[0.23]	[0.04]	[1.00]	[0.58]		
Weeks worked last year	44.79	0.153	0.168	-0.0015	2.46	1.013	2.3%
	(6.031)	(0.0780)	(0.0637)	(0.0078)	(0.12)	[.002,2.221]	
		[0.25]	[0.04]	[1.00]	[0.58]		
Usual hours works per week	41.31	0.148	0.1722	0.0039	2.94	0.98	2.4%
	(6.040)	(0.0882)	(0.0798)	(0.0080)	(0.09)	[-.161,2.317]	
		[0.37]	[0.12]	[1.00]	[0.52]		
Log labor income	10.88	0.0033	0.00238	-0.0008	0.33	0.0219	
	(0.301)	(0.0043)	(0.0036)	(0.0005)	(0.57)	[-.035,.083]	
		[1.00]	[1.00]	[0.61]	[1.00]		
Log family income/poverty	5.91	0.0035	0.00245	-0.0006	0.07	0.023	
	(0.302)	(0.0048)	(0.0041)	(0.0005)	(0.79)	[-.04,.092]	
		[1.00]	[1.00]	[1.00]	[1.00]		
In poverty*	0.07	-0.00033	-0.000637	-0.0001	0.01	-0.002	-2.9%
	(0.095)	(0.0013)	(0.0011)	(0.0001)	(0.90)	[-.02,.016]	
		[1.00]	[1.00]	[1.00]	[1.00]		
Rec'd public assistance*	0.11	-0.0046	-0.0046	0.0002	9.93	-0.029	-26.4%
	(0.110)	(0.0015)	(0.0011)	(0.0001)	(0.00)	[-.056,-.01]	
		[0.02]	[0.00]	[1.00]	[0.01]		
B. Women							
Self-sufficiency index	0.019	0.0058	0.00406	0.0005	4.24	0.0399	
	(0.18)	0.0029	0.0027	0.0004	[0.10]	[.001,.088]	
Worked last year	0.8	-0.0018	-0.00168	0.0001	0.15	-0.012	-1.5%
	(0.13)	(0.0020)	(0.0018)	(0.0003)	(0.70)	[-.043,.016]	
		[0.90]	[1.00]	[1.00]	[1.00]		
Weeks worked last year	37.66	-0.097	-0.0931	0.0091	0.03	-0.667	-1.8%
	(6.80)	(0.1100)	(0.1015)	(0.0146)	(0.87)	[-2.31,.858]	
		[0.90]	[1.00]	[1.00]	[1.00]		
Usual hours works per week	30.49	-0.11	-0.092	0.0117	0.00	-0.745	-2.4%
	(5.81)	(0.0970)	(0.0950)	(0.0129)	(0.95)	[-2.24,.599]	
		[0.90]	[1.00]	[1.00]	[1.00]		
Log labor income	10.31	0.0062	0.007	0.00084	3.94	0.043	
	(0.34)	(0.0051)	(0.0048)	(0.0005)	(0.04)	[-.026,.122]	
		[0.90]	[0.57]	[0.61]	[0.28]		
Log family income/poverty	5.81	0.0083	0.0084	-0.00034	0.69	0.057	
	(0.32)	(0.0054)	(0.0050)	(0.0005)	(0.41)	[-.015,.143]	
		[0.62]	[0.48]	[1.00]	[1.00]		
In poverty*	0.12	-0.0028	-0.00273	-0.0003	4.21	-0.018	-15.0%
	(0.11)	(0.0017)	(0.0015)	(0.0002)	(0.04)	[-.047,.004]	
		[0.60]	[0.38]	[0.71]	[0.28]		
Rec'd public assistance*	0.12	-0.0049	-0.0029	0.0002	0.67	-0.034	-28.3%
	(0.12)	(0.0014)	(0.0011)	(0.0001)	(0.41)	[-.06,-.014]	
		[0.00]	[0.06]	[0.71]	[1.00]		

Notes: *In poverty and received public program income are reverse-coded when used in the self-sufficiency index. See also Table 2.1 notes.

Table 2.5. Heterogeneity in the Effect of Head Start, by Local Programs and Economic Circumstances

	Intent-to-Treat (ITT) Effect			Effect on preschool enrollment (s.e.)		Average Treatment Effect on Treated Children (ATET)	
	Above median	Below median	F-test of diff (p-val)	Above median	Below median	Above median	Below median
A. Human capital index							
Medicaid exposure	0.021	0.009	4.095	0.123	0.175	0.171	0.052
	(0.004)	(0.004)	(0.043)	(0.024)	(0.030)		
CHC exposure	0.020	0.010	9.258	0.203	0.109	0.100	0.090
	(0.003)	(0.003)	(0.002)	(0.034)	(0.025)		
Food Stamps exposure	0.011	0.020	4.515	0.122	0.167	0.092	0.117
	(0.003)	(0.004)	(0.034)	(0.026)	(0.032)		
Poorest 300th counties	0.009	0.015	0.320	0.073	0.125	0.125	0.123
	(0.011)	(0.003)	(0.572)	(0.026)	(0.023)		
Predicted economic growth	0.023	0.010	11.948	0.139	0.125	0.166	0.078
	(0.004)	(0.003)	(0.001)	(0.033)	(0.025)		
B. Self-sufficiency index							
Medicaid exposure	0.003	0.005	0.336	0.123	0.175	0.023	0.031
	(0.004)	(0.002)	(0.562)	(0.024)	(0.030)		
CHC exposure	0.006	0.003	1.708	0.203	0.109	0.030	0.028
	(0.003)	(0.003)	(0.191)	(0.034)	(0.025)		
Food Stamps exposure	0.002	0.007	3.547	0.122	0.167	0.017	0.039
	(0.003)	(0.002)	(0.060)	(0.026)	(0.032)		
Poorest 300th counties	0.004	0.004	0.000	0.073	0.125	0.060	0.035
	(0.010)	(0.002)	(0.991)	(0.026)	(0.023)		
Predicted economic growth	0.006	0.003	1.014	0.139	0.125	0.041	0.027
	(0.003)	(0.003)	(0.314)	(0.033)	(0.025)		

Notes: ATETs are constructed by dividing the group-specific ITT estimate of Head Start's effect on long-run outcomes by the group-specific estimated first stage. Results for the 300 poorest counties are reported in the column for "above median" with results for other counties reported in "below median."

CHAPTER III

Birth Order and Unwanted Fertility

Abstract

An extensive literature documents the effects of birth order on various individual outcomes, with later-born children faring worse than their siblings. However, the potential mechanisms behind these effects remain poorly understood. This paper leverages U.S. data on pregnancy intention to study the role of unwanted fertility in the observed birth order patterns. We document that children higher in the birth order are much more likely to be unwanted, in the sense that they were conceived at a time when the family was not planning to have additional children. Being an unwanted child is associated with negative life-cycle outcomes as it implies a disruption in parental plans for optimal human capital investment. We show that the increasing prevalence of unwantedness across birth order explains a substantial part of the documented birth order effects in education and employment. Consistent with this mechanism, we document no birth order effects in families who have more control over their own fertility.

JEL Codes: J13, J22, J24

3.1 Introduction

A large literature in psychology *and* economics documents clear patterns of birth order effects on various human capital and labor market outcomes. However, little is known about the potential underlying mechanisms behind these effects. Evidence suggests that later-born children receive less parental time and investment compared with their earlier-born siblings, but the sources of these differences remain poorly understood.

This paper contributes to the literature on birth order effects by proposing a novel explanation for a significant part of the observed patterns. It connects the literature on birth order effects and the literature on the negative impacts of unwanted fertility over the life cycle.[1] We show that the observed birth order effects on education and employment may reflect a disruption in parental plans for optimal investment in their children, as later-born children are more likely to have been conceived when the family did not want to have an additional child.

We believe the complementary mechanism we propose to be novel as it provides distinct testable implication relative to current explanations for observed birth order effects. In particular, as we discuss below, standard theories predict that birth order effects would apply to all families irrespective of the pregnancy intention statuses of the various children in the birth order sequence. Our proposed mechanism implies that the existence or at least the size of a birth order gradient may depend on the fertility control capabilities of various families.

It is natural to expect a higher incidence of excess, unwanted children at higher parities. Few families who are still childless are likely to report that they were not planning to have children, so very few first-born children are likely to be unwanted. Some families may want

[1] We follow the demographic definition of an *unwanted* birth as a birth in excess of total desired fertility. The broader notion of *unintended* birth includes both unwanted and mistimed births. Given our purposes, we do not consider mistimed births.

to have only one child, so second-born children are more likely to be unwanted. Most families desire to have only two children, so it is likely that the incidence of unwanted children among third-borns is much higher.

The very occurrence of an unwanted birth may give rise to birth order differences. While most families may wish to equalize human capital investments among their children, only families that succeed in realizing their fertility plans are able to do so. For example, those who only plan to have two children may make irreversible decisions that might only be consistent with equalization between the first two (wanted) children. Once a third (unwanted) child is born, the family's original investment plan is disrupted and it becomes ex-post suboptimal. While some re-optimization could occur, particularly by reducing planned investments or disinvesting in first- and second-born children, it might be very costly or impossible at that point to re-optimize in a way that achieves perfect equalization among all three siblings. This friction could give rise to the birth order effects we see in the data. There might be rigidities to disinvesting in existing siblings and those disinvestments might be necessary to reach perfect equalization. Moreover, the rigidity might be larger for the oldest, first-born sibling, since more time has passed and more investments have been embodied and crystalized in that child.

According to this story, if parents have made commitments that are difficult to revise upon the birth of an unwanted child, then the unwanted, higher birth order child will tend to receive less investment relative to earlier born siblings, generating a gap that contributes to the birth order effects. For example, parents may have made location, labor supply, housing, consumption, borrowing, and investment decisions that would have been different had they known an additional child would be eventually born. As a result, higher order, unanticipated children may tend to do worse in terms of human capital and labor market outcomes, and this could contribute to the birth order gradient. A corollary is that birth order effects would tend to be ameliorated in families that for various reasons have more control over their own fertility.

In this paper, we first replicate earlier findings of birth order effects using both OLS and family fixed effects specifications. Next, we show that these effects vanish once we focus on a subsample of families who intended to have all the children they had. We then show that the incidence of unwanted births increases significantly with birth order, and that accounting for this pattern of unwantedness flattens the birth order gradient. We also show that our birth order results are robust to alternative mechanisms, such as exposure to changes in family structure or last-born effects arising from endogenous fertility stopping rules. When we investigate subgroups, we find that birth order effects no longer arise when we restrict our focus to families that presumably have (for religious reasons) more control over their fertility.

The data requirements to accomplish this study are somewhat demanding. We rely on longitudinal microdata from the U.S. Panel Study of Income Dynamics. The PSID data allow us to observe the adult outcomes of multiple siblings within a family along with their birth order. Of more novelty and critical for our purposes, the PSID data includes a retrospective maternal assessment about her pregnancy intention status at the time each of these siblings were conceived. We argue this information is particularly valuable for any study that aims to understand the role of birth order. Without it, one might be missing a significant element of parental decision-making.

The rest of the paper proceeds as follows. In Section 3.2, we discuss the two unrelated literatures that converge in this paper. In Section 3.3, we describe the PSID data we use in this paper. Section 3.4 presents our empirical methods and main findings. Section 3.5 provides concluding remarks.

3.2 Related Literature

In the two subsections below, we provide a brief overview of relevant previous work in the two strands of literature that have so far developed independently of each other. First, we summarize a large literature that documents and attempts to explain birth order effects in various outcomes. Second, we summarize a smaller literature that explores the association between maternal pregnancy intention and individual outcomes later in life. To our knowledge, there is no previous work connecting these two literatures explicitly as we propose in this article.

3.2.1 Birth Order

A vast literature in economics explores birth order effects in various outcomes over the life cycle. Much of the literature focuses on completed education as the outcome of interest, and uses data from developed countries finding large and significant negative effects associated with a higher birth order.[2] The earliest work on birth order we are aware of in the economics literature goes back to Lindert (1977). He pointed out the importance of exploiting within family variation to ensure that no unobserved family characteristics confound the birth order patterns. He suggested that the family's time budget gets diluted across various siblings and this process tends to benefit earlier-born siblings. A modern testing of this hypothesis can be found in the influential work by Price (2008). Also seminal to this literature was work by Behrman and Taubman (1986), who explored birth order effects within the theoretical framework of Behrman, Pollak and Taubman (1982).[3]

The literature was re-invigorated in the 2000s with the work of Black, Devereux and Sal-

[2]The birth order literature is more limited in developing countries so little is known about how these patterns generalize to that context. For a few exceptions, see Birdsall (1979), Birdsall (1990), Behrman (1988), Horton (1988), Ejrnæs and Pörtner (2004), Edmonds (2006) and De Haan, Plug and Rosero (2014)

[3]See Kessler (1991) for additional early references.

vanes (2005), who documented convincing birth order effects using a large dataset from Norway and spurred a large and still active literature in economics. Subsequent work has documented birth order effects on various outcomes, primarily focusing on school attainment and performance (Kantarevic and Mechoulan (2006), Booth and Kee (2009), De Haan (2010), Bagger et al. (2013), Hotz and Pantano (2015)), measures of cognitive ability (Conley and Glauber (2006), Black, Devereux and Salvanes (2011), Lehmann, Nuevo-Chiquero and Vidal-Fernandez (2013), Hotz and Pantano (2015), Pavan (2016)), but also exploring risky behaviors (Argys et al. (2006), Hao, Hotz and Jin (2008), Averett, Argys and Rees (2011)), health (Black, Devereux and Salvanes (2016), Björkegren and Svaleryd (2017), Brenøe and Molitor (2018)) and delinquency (Breining et al. (2017)).

Much of the recent literature has tried to uncover and elucidate various alternative mechanisms giving rise to observed birth order effects that go beyond the above-mentioned "time dilution" theory. These hypotheses range from differential parenting strategies (Hao, Hotz and Jin (2008), Averett, Argys and Rees (2011), Hotz and Pantano (2015)) to parental transfer behavior (De Haan (2010), Mechoulan and Wolff (2015)), direct parental preferences for birth order (Bagger et al. (2013)) and early parental investments (Lehmann, Nuevo-Chiquero and Vidal-Fernandez (2013), Pavan (2016)). Birth order effects have of course been studied in other disciplines, primarily in psychology. Early work by Belmont and Marolla (1973) provided some of the more convincing empirical evidence that spurred most of the modern research on birth order in recent decades, but some of these arguments go back to Galton (1875). Sulloway (2010) and Eckstein et al. (2010) provide surveys of the literature in psychology which focus more on how personality and non-cognitive skills vary with birth order. This is an important outcome that economists are beginning to tackle as well (Black, Grönqvist and Öckert (2017)).

Overall, this literature has provided substantial evidence that earlier-born siblings tend to have higher cognitive skills and very different personality traits, go on to complete higher levels of schooling, engage less frequently in risky behaviors, earn higher wages, and display a variety of positive outcomes along various dimensions later in life.

3.2.2 Unwanted Fertility and its Effects

The standard model of completed fertility and child quality in economics (Becker and Lewis (1973), Willis (1973) and much of the large literature that followed these seminal contributions) assumes that fertility is a perfectly controlled process, so unwanted births are not defined in that setting. Yet a growing literature, primarily in demography, has documented the prevalence of unwanted pregnancies and the life-cycle consequences associated with being the result of an unwanted pregnancy.[4] Michael and Willis (1976) extended these early economic models to incorporate imperfect fertility control and a distinction between desired and realized births. Economists have also recognized early on the importance of allowing for the effects of a dynamically changing, uncertain environment in the econometric modeling of reproductive decisions (see, for example, Barmby and Cigno (1990)).

Economists have also begun to investigate these issues empirically by exploiting direct maternal assessments of pregnancy intention (Rosenzweig and Wolpin (1993), Joyce, Kaestner and Korenman (2000), Joyce, Kaestner and Korenman (2002), Miller (2009), Lin and Pantano (2015), Lin et al. (2017)). Others have exploited natural experiments or reproductive policy changes that allow comparison of individuals from otherwise similar cohorts that were born at times when mothers had greatly different opportunities to control their own fertility (Gruber, Levine and Staiger (1999), Donohue and Levitt (2001), Charles and Stephens (2006), Pop-Eleches (2006), Donohue, Grogger and Levitt (2009), Ananat and Hungerman (2012), Ozbeklik (2014)) or at times that are thought to be auspicious for birth (Do and Phung (2010)). The balance of the literature indicates that unwanted children tend to do worse along many dimensions of adult life (education, employment, health, crime, etc.)

[4] See among others Baydar (1995), Kubička et al. (1995), Myhrman et al. (1995).

3.3 The Data

Our primary source of data is the Panel Study of Income Dynamics (PSID), a longitudinal survey of a nationally representative sample of US individuals and families with ongoing data collection since 1968. The PSID continuously collects information on individual longitudinal outcomes for its initial survey respondents and their descendants. We can observe individuals and their siblings. This allows us to study the effects of birth order on outcomes later in life (e.g., completed education, employment) by comparing children of different birth order within the same families.

We construct our sample of children from the Childbirth and Adoption History File. The file contains records of childbirths and adoptions of individuals living in a PSID family at the time of the interview in any wave from 1985 through 2013. We restrict our sample to childbirth events reported by mothers and drop multiple births. We then obtain details about each child, including year of birth and birth order, as well as details about the mothers, such as their age at the time of birth, total number of births, and year of the most recent maternal report. We only include in our sample the children with mothers whose most recent report happens after she turns 35, so her total number of births is unlikely to change in the future.

Using the unique identifiers in the PSID, we combine the childbirth information with other PSID files at the individual and family levels to construct our set of control variables and to add key pieces of information for our study. First, we construct completed years of education for our sample of children. Using the PSID cross-year individual file, we only collect years of education after age 24, or when the same individual's highest years of education appear in two or more waves and consistent with the latest record, so the process of human capital accumulation is most likely to have completed. We also construct a measure of employment in adulthood. The employment variable is constructed by reverse-coding an indicator of whether the individual was unemployed at any time in 2011 and restricted to individuals who were between the ages of 24 and 50.[5]

We also obtain valuable information on pregnancy intention reported by the mothers of our sample children in the PSID. In the 1985 interview, the PSID included a questionnaire for wives and long-term female cohabiters, allowing these women to answer for themselves some questions about their fertility history. Included in the set of questions unique to the 1985 interview is a retrospective pregnancy intention assessment. We use these pregnancy intention reports to construct our indicator defining a child as unwanted. Specifically, we define a child as unwanted if the mother reported that she was not planning to have any (more) children when she became pregnant with that child. This means that neither the children whose mother reported them as "mistimed" nor those reported as "wanted" will be considered unwanted. In particular, mistimed children who were conceived "too soon," while still the result of unintended pregnancies, are not considered *unwanted* according to our definition. Parents had a plan to eventually have this child, and their decision making was conditional on that plan.[6]

Our analysis is limited by the fact that pregnancy intention questions were only fielded in the 1985 survey. We are therefore unable to evaluate pregnancy intention status for individuals born after 1985. However, this limitation is not significantly stricter than the requirement for our sample children to have completed their education by 2013. It is also worth noting that the mothers were only asked to report pregnancy intention associated with the conception of their first, last, and second to last child, which means we can only observe the complete pattern of pregnancy intentions for families with no more than three children. This limits our ability to look at larger families. Still, the available data allow us to examine birth order effects in families with two or three children. This will be sufficient to convey our main

[5]The indicator is not defined for those who reported being out of labor force for the entire year.

[6]Children who result from mistimed pregnancies, particularly when these occur before marriage, may also have negative effects on outcomes later in life. See, for example, Nguyen (2018)

findings.

The use of retrospective assessments of pregnancy intention is controversial, as many (see, for example, Westoff and Ryder (1977) and Rosenzweig and Wolpin (1993)) fear that these reports could be contaminated by ex-post rationalization and other selective recall problems. However, work by Schoen et al. (1999) and Joyce, Kaestner and Korenman (2002) specifically addresses these questions and suggests that these retrospective reports about pregnancy intention tend to be valid.[7] We present summary statistics for our analysis sample in Table 3.1.

As can be seen in the table, our sample comes from families with two or three children so the incidence of third-born children in the sample is smaller than that of either first-born or second-born children.[8] The average number of years of completed education in the sample is about 14. The sample is primarily composed of white and black children with a small number of Hispanics and children of other races. The average age of the children in our sample is 40 as of 2013. By 2013, the children in our sample have all grown up and are all well into their adult years. About 16 percent of these children are unwanted as defined above. However, as we will now show, this masks substantial variation across birth order.

The pattern of unwantedness by birth order for subsamples of children from families with two or three children is presented in Table 3.2. In families with two children, 11% of first-borns are unwanted. The rate of unwantedness increases to 15% for second-born children.

Families with three children show the same pattern of increasing prevalence of unwanted

[7] Joyce, Kaestner and Korenman (2002) find that prospective and retrospective reports of pregnancy intention provide the same estimate of the effects of being an unintended child on various prenatal outcomes once they control for selective pregnancy recognition using an IV procedure. Further, they show that the extent of unwanted fertility was the same regardless of whether the assessment was during pregnancy or after birth. They show this for a subsample of women for whom pregnancy intention was assessed both prospectively (during pregnancy) and retrospectively (after birth).

[8] In principle, since we are looking at families with two and three children, the number of first-born and second-born children should be the same. In practice however, our number of second-born children is slightly smaller than the number of first-borns because they are more likely to have missing information on our outcome of interest, completed education.

pregnancy with birth order. Notably though, 37% of third-born children are reported as unwanted by their mothers, a substantial increase relative to first and second birth order. As discussed earlier, we are unable to document unwantedness status for all children in families of more than three children. In addition, a much smaller number of families in PSID has four or more children. For these two reasons, we limit our analysis to families with two or three children. [9]

It is of interest to explore whether the pattern we identify for our PSID children born before 1985 holds also for more recent cohorts. Families who had all of their children more recently may have been in better position to plan their fertility and not exceed their desired family size. To explore this, we rely on data from various recent waves of the National Survey of Family Growth (NSFG) as reported in Child Trends (2013). The NSFG is a repeated cross-sectional survey and, as such, cannot be used to explore birth order effects in adult outcomes like completed schooling. However, the NSFG includes a valuable retrospective assessment of a woman's fertility history that can be used to explore how pregnancy intention varies with birth order. Table 3.3 combines information from our PSID sample with reports based on the NSFG.

As can be seen in the table, there is still a substantial increase in the prevalence of unwantedness as we move across the birth order in more recent years, particularly for those who are third-born or have an even higher birth order. The incidence of unwanted children among third-borns is somewhat smaller in the more recent cohorts that can be reported about in the NSFG waves. This is because our PSID sample of children draws from earlier cohorts where the opportunities to prevent unwanted births were more limited. In particular, at the time in which many of the children in these earlier cohorts were conceived, abortion

[9] However, it is of interest to explore whether the pattern of increasing prevalence of unwanted children across birth order holds in 4-child families. Since we only know whether the first, last, or second to last child in a family was unwanted, we cannot tell whether a second-born child in a four-child family was unwanted. But we can still look at first-, third-, and fourth-born children in those families. Consistent with the patterns in Table 3.2, we find that in four-child families, the incidence of unwanted children grows from 16% among first-borns, to 27% among third-borns to a whopping 53% among fourth-borns.

was not yet legal, and oral contraceptives were not yet widely available. Lin and Pantano (2015) document a decrease in the prevalence of unintended births following legalization of abortion. Bailey (2010) points out the large increase in the share of families with fewer than three children following the "contraceptive revolution" of the 1960s. Presumably much of that change was accomplished by the avoidance of what would have otherwise been unwanted *third-born* children. However, as the NSFG numbers show, while in the more recent years the opportunities to prevent or terminate unwanted pregnancies are more widely available, it remains the case that the prevalence of unwanted births increases with birth order and it is particularly high for third- and higher-born children.

3.4 Methods and Results

To examine the relationship between birth order and completed years of schooling, we consider the following model in the same mold as Kantarevic and Mechoulan (2006) and others in the birth order literature:

$$Y_{ih} = \alpha_1 + \alpha_2 1\left[BO_{ih} = 2\right] + \alpha_3 1\left[BO_{ih} = 3\right] + \beta X_{ih} + \varepsilon_i \tag{3.1}$$

where Y_{ih} denotes completed years of education of child i in family h,$1\left[BO_{ih} = k\right]$ is an indicator variable that equals 1 whenever child i has the kth birth order in family h. X_{ih} is a vector of control variables that accounts for observed characteristics of child i and/or family h. Note that first-borns correspond to the omitted category.

We begin by replicating the results reported in Kantarevic and Mechoulan (2006), who also use the PSID data. Table 3.4 presents our results. We are successful at replicating their main findings. The results are presented along six columns.

The first three specifications do not include controls for X_{hi}, whereas the last three do. In both cases, the first specification, in columns (1) and (4), pools families with 2 and 3 children (controlling for family size) and the following two specifications consider models separately for a subsample of families with 2 children and a subsample of families with 3 children. Standard errors are clustered at the family level in all of our specifications.

As can be seen in column 1, in the simplest specification without controls, second- and third-born children tend to complete 0.09 and 0.25 fewer years of education, respectively. Once we control for the child's sex and race as well as family size and maternal characteristics such as mother's age at birth and mother's education, we find that the birth order effects are accentuated. The negative coefficients on the indicators for second and third born children are now larger in magnitude (-0.31 and -0.44) and more statistically significant. Estimates for control variables in columns 4–6 have the expected sign: males and children in black families tend to have completed less schooling. Family size has overall a negative and statistically significant effect, with those growing up in families with three children having on average 0.20 fewer years of schooling than those in 2-child families. Controlling for family size is particularly important when pooling children from families of different sizes because, by construction, a higher birth order is only feasible in larger families.

While some observed characteristics of the mother and the family can be controlled for, it is always possible that there are additional unobservable characteristics that could confound the effects of birth order. To tackle this issue, we exploit information on siblings of different birth order within the same families. To do so we follow Kantarevic and Mechoulan (2006) and explore a family fixed effects specification:

$$Y_{ih} = \alpha_1 + \alpha_2 1\left[BO_{ih} = 2\right] + \alpha_3 1\left[BO_{ih} = 3\right] + \beta X_i + \lambda_h + \varepsilon_{ih} \tag{3.2}$$

where the only differences with respect to the model in (3.1) is that X_i now is a vector of control variables that only accounts for observed characteristics of child i within a family

h as the family characteristics, both observed and unobserved are absorbed into the family fixed effect λ_h. Results of estimating the model in (3.2) are presented in Table 3.5.

They show that the birth order patterns are robust to controlling for family characteristics that are common across siblings within a given family. Controlling for family fixed effects and thus using only within family variation produces estimated birth order effects that remain sizable and significant, especially among three-child families. In particular, the pooled specification still shows significant reductions of 0.24 and 0.39 years of schooling for second- and third-born children relative to their own first-born siblings.

3.4.1 Imperfect Fertility Control

Our main hypothesis is that families who have the ability to perfectly control their fertility are better able to equalize outcomes among their offspring. We conjecture that families with more imperfect fertility control are less likely to accomplish such equalization. This could be because unwanted children in excess of the family's target level of desired fertility are not "budgeted for" in the family's investment plan. It is possible for the family to re-optimize upon the birth of an unwanted child by re-allocating resources from elder siblings to the new unwanted child. But such re-allocations are unlikely to accomplish a perfect equalization.

To investigate this possibility, we next explore how results in Table 3.5 change when we focus on families for which there is evidence of perfect fertility control and families for which there is not. To implement this, we create an indicator $Unwanted_{hi}$ which equals one whenever child i in family h was the result of a pregnancy that was retrospectively assessed as unwanted by the mother. A child is defined as unwanted in the sense described in Section 3.3. We then define $W_h = 1$ if all children in family h are reported as wanted, and we set $W_h = 0$ otherwise. Panels A and B of Table 3.6 present the results in the two subsamples ($W_h = 1$ and $W_h = 0$).

We first re-estimate the family fixed effects specification in (3.2) for a subsample of families where all children are reported as wanted ($W_h = 1$). These are families who planned to have all of the children they ended up having, and therefore faced no disruption in their optimal child investment allocation process. These families represent 50% of all the families whose children we use in the full sample. Panel A of Table 3.6 shows the results.

As can be seen in the table, the birth order effects we documented in Table 3.5 essentially disappear. The coefficients on second- and third-born children are now much smaller and, in some cases, even positive for these families who had no unwanted children. Moreover, none of the birth order effects in panel A are statistically significant.

On the other hand, panel B of Table 3.6 shows that families without evidence of perfect fertility control ($W_h = 0$) retain sizable and significant birth order effects.[10]

Panel C presents tests of statistical significance for the differences between panels A and B. In the pooled specification, the differences are statistically significant at 1% for the second child and 5% for the third child. The differences for the second child are also significant even in the smaller subsamples of 2- and 3-children families. These substantial differences provide our first line of evidence showing that birth order effects are somehow linked to families' fertility control.

While the differences in the birth order gradient are striking, we exercise caution when interpreting these results. It is possible that unobserved family heterogeneity could be driving our findings. That would be the case if some family unobservable is correlated with both, imperfect fertility control and the existence of birth order effects. Families who are good at avoiding unwanted births are different (in observed and likely unobserved ways) from other families who are less successful regarding fertility planning. Therefore, we do not claim to attach a causal interpretation to these results. For example, it could be that families good

[10] These are families for which we identify at least one unwanted child or families for which information for pregnancy status is missing for at least one child.

at contracepting are particularly averse to inequality in outcomes and thus try harder to avoid a large birth order gradient among their offspring. If that is the case, one would see more inequality in outcomes among the children of families who contracept poorly. This is something to bear in mind as the family fixed effects in our child level models do not necessarily account for this problem.

3.4.2 Accounting for Pregnancy Intention in Estimation of Birth Order Effects

The family fixed effects specification has become standard when looking at birth order effects, but it only controls for unobserved factors that are common across siblings within a family. We have already controlled for maternal age at birth, but there might be other characteristics that are typically unobserved, vary across siblings, and are correlated with both birth order and later life outcomes like completed years of schooling. One such factor we do get to observe is the pregnancy intention status corresponding to the conception of each child in the family. As documented in Section 3.3, higher birth order children are more likely to be the result of an unwanted pregnancy. Given the large differences observed in panels A and B of Table 3.6, it is natural to explore how birth order effects change once we account for this child-specific factor, often unobserved in various datasets that are used to document birth order effects.

Before exploring how accounting for pregnancy intention affects birth order effects on completed education in the full sample, we provide a more systematic examination of how the chance of being unwanted rises with birth order. While the results in Table 3.2 are quite suggestive, we first investigate whether those results are robust to controlling for the same set of observable characteristics X. To do so, we re-estimate the model in (3.1) using our indicator that denotes whether the individual was the result of an unwanted pregnancy as dependent variable. The results are presented in Table 3.7.

As we can see, the pattern of increasing unwantedness as we move higher in the birth order is robust to controlling for observable characteristics. Moreover, as can be seen in Table 3.8, the results are robust to further controlling for family fixed effects, particularly for third-born children.

Having established a clear pattern of increasing prevalence of unwanted children higher on the birth sequence, we explore the implications of pregnancy intention across birth order for the estimation of birth order effects in education. To that end, we include our key measure characterizing each child within a family as wanted or unwanted in our models for completed education. Our objective is to explore if and how the pattern of birth order effects on education changes once we control for the maternal pregnancy intention indicators corresponding to each child. Table 3.9 presents results from family fixed effects specifications that add the child-level $Unwanted_{hi}$ indicator to the model in (3.2).

The results show that the birth order gradient becomes less pronounced once we account for indicators characterizing maternal pregnancy intention at the time these children were conceived.

For example, in the pooled specification in column 1, the effects for second- and third-born children change from -0.24 and -0.39 in Table 3.5 to -0.14 and -0.23 in Table 3.9, both large percent-wise reductions in magnitude. Moreover, using a 1000-resamplings bootstrap, we found that the birth order effect differences in the pooled sample of two- and three-child families across Tables 3.5 and 3.9 are both statistically significant at the 10% level. Further, we now see that only higher birth order children in three-child families have a significantly negative effect, but the magnitudes (-0.25 for second-born and -0.37 for third-born) are much smaller than those in Table 3.5 (-0.30 for second-born and -0.50 for third-born). Indeed, for third-born children in three-child families, the difference across tables is, again, statistically significant at the 10% level.

Our focus here is not to estimate the impact of being an unwanted child, but rather account for differential pregnancy intention rates across birth order in the estimation of birth order effects. Further, we aim to explore how different the birth order gradient is across families with and without imperfect fertility control. However, it is worth noting that the $Unwanted_{hi}$ indicator is not statistically different from zero in the family fixed effect specifications of Table 3.9. This is consistent with work by Joyce, Kaestner and Korenman (2000) who find similar results using the National Longitudinal Survey of Youth, albeit looking at outcomes earlier in life. This is not surprising because, as pointed out by Rosenzweig and Wolpin (1993), the family fixed effects specification will tend to provide a biased estimate of the effects of being unwanted. This is because the birth of an unwanted child will likely have an effect on existing siblings as parents re-optimize (i.e., reduce) their allocations toward those siblings as they cope with the arrival of the unwanted child. While our main approach takes as given the idea that unwanted children do worse, neither our OLS (because of omitted variable bias) nor our family fixed effect (because of spillover on existing siblings) strategies are well suited to test that hypothesis in a causal sense. The true effect is probably somewhere in between.

Taken together, the attenuation of the birth order gradient once we account for unwanted births, coupled with the absence of birth order effects in families where all children are wanted, suggests that pregnancy intention might be an important consideration when assessing the effects of birth order on various outcomes.

3.4.3 Birth Order Effect Heterogeneity in Groups with More or Less Imperfect Fertility Control

In this subsection, we explore how birth order effects vary among groups with differential fertility control. We follow Lin and Pantano (2015) and use information on maternal religion to classify our children into two groups. First, we create a group whose mothers report a religion affiliation that tends to be more strongly against the use of abortion. We denote this as the "pro-life" subsample. We group a second set of children whose mothers report either no religion affiliation or an affiliation that has less stringent attitudes towards abortion. We denote this as the "pro-choice" subsample. We use the same criteria as in Lin and Pantano (2015) to classify these religions into this binary indicator of attitudes toward abortion.[11] To be sure, within each religion, there will be those who adhere more strictly to their religion's stance and those who will align less strongly. The indicator is not meant to classify the exact attitude of a particular mother, but rather provide an indicator of her *likely* ability to terminate unwanted pregnancies. We expect the birth order effects to be stronger in the "pro-life" sub-sample as mothers in this sub-sample are less likely to use abortion to terminate unwanted pregnancies.

On the other hand, we expect the birth order effects to be milder in the "pro-choice" sub-sample as these families are more likely to use abortion to prevent unwanted births. As a result, the prevalence of unwanted third-born children, relative to first-born children may not be as high for these "pro-choice" families. We re-estimate the family fixed effects specification in (3.2) in the sub-samples of children grouped according to these different maternal religious affiliations. Tables 3.10 and 3.11 present the results.

The birth order effects are quite strong in the "pro-life" subsample. Indeed, they are as strong as those reported in Table 3.5. For example, in the pooled specification, on average, second- and third-born siblings complete 0.26 and 0.40 fewer years of education than their first-born sibling within the same family.

On the other hand, there are no apparent birth order effects in the "pro-choice" subsample

[11] We follow the religion taxonomy in Evans (2002) and classify the following religions as having a more strict attitude against abortion: Roman Catholic, Protestant, other Protestant, other Non-Christian, Latter Day Saints, Mormon, Jehovah's Witnesses, Greek/Russian/Eastern Orthodox, Lutheran, Christian, Christian Science, Seventh Day Adventist, Pentecostal, Jewish, Amish, and Mennonite. Mothers reporting these religions are more likely to be pro-life and less likely to use abortion to terminate unwanted pregnancies. We then classify Baptists, Episcopalians, Methodists, Presbyterians, and Unitarians along with Agnostics and Atheists as having a less strict attitude towards abortion.

as reported in Table 3.11. As can be seen in this table, while the point estimates are still mostly negative, all of them are smaller in magnitude and none of them is statistically significant. However, while the differences in the birth order estimates across the two groups are sizable, we are unable to reject the hypothesis that the effects in Tables 3.10 and 3.11 are the same.

3.4.4 Alternative Mechanisms

We have found compelling evidence of the important role played by pregnancy intention in generating birth order effects. We now test two alternative mechanisms that could give rise to birth order effects in our sample.

First, it has been documented that family structure may affect children's education outcomes (Ermisch and Francesconi (2001)). Then, it is possible that children higher in the birth sequence could be more affected by changes in family structure relative to earlier-born siblings and this could affect their educational achievement. As a result, it has become standard in the literature to test whether birth order effects hold in a subsample of intact families (see, for example, Black, Devereux and Salvanes (2005), Hotz and Pantano (2015)). Second, it is possible that endogenous fertility-stopping rules could be the reason behind our birth order effects. We show that the birth order effects we identify in this sample are robust to these two possibilities. To test whether our results in Table 3.5 reflect just differential exposure to changes in family structure, we re-estimate our main fixed effects specification in (3.2) but in a subsample of intact families.

To construct our subsample of intact families, we link our data with the PSID Marriage History File, which contains information on the mothers' marriage events collected retrospectively in the 1985 through 2013 waves. We use this data to create a sample of intact families within our main sample by keeping a family only when the mother's first marriage stayed intact by the year her last child turned 24, and when the mother's most recent marriage report was collected after her last child turned 24.

By applying these sample restrictions, we focus on a subsample of individuals whose family structure was relatively stable through the completion of their education. Results are presented in Table 3.12.

As we can see in the table, sizable birth order effects are still present in this subsample, and remain statistically significant despite the smaller sample size. This suggest that the birth order effects we find are not driven by differential exposure to disruptions in family structure. Later-born siblings have lower educational attainment for reasons other than their higher likelihood of growing up in a broken home.

Next, we investigate whether the birth order effects we document are driven by endogenous fertility stoppage. When a newborn child is particularly unhealthy, it is possible that parents may not have additional children, to ensure they have the time and resources to care for the unhealthy child. If this is the case, the last born would tend to be particularly unhealthy and this could affect cognitive development and, ultimately, educational attainment. Note that later-born siblings tend to engage more in risky behaviors in their late teen and early twenties and this could lead to reduced health in adulthood, but would not necessarily provide evidence consistent with the rationale behind an endogenous fertility rule. Still, in our preliminary explorations, we found no birth order effects in measures of adult health. To investigate this further, we deemed more appropriate to use a measure of health earlier in life. We re-estimated our main fixed effects specification in (3.2) but using as dependent variable a measure of health during childhood and adolescence.

This presents a challenge as the individuals in our sample grew up during years in which the PSID was not yet systematically collecting information on health. Fortunately, in more recent years, PSID asked heads of households and wives to provide a summary retrospective assessment about their own health status earlier in their lives (health status before they turned 17 years old). This measure could be more plausibly related to the type of health issues that could lead an individual's parents to stop their fertility. Results are presented in Table 3.13. As can be seen in this table, we find no significant birth order effects in health status during childhood and adolescence. These findings ameliorate concerns that our results could be driven by endogenous last-born effects.

3.4.5 Employment Outcomes

In this subsection we explore how birth order is associated with adult employment and whether pregnancy intention plays a similar role. The construction of our measure of employment in adulthood is described in Section 3.3. Our findings are broadly similar to the reported effects on completed education.

Tables C.1-C.7 in the Appendix present the results. Table C.1 presents the basic OLS estimates and Table C.2 reports the basic family fixed effects estimates. These tables show that there is a reduction in the probability of adult employment for second- and third-born children relative to first-born children. In the family fixed effects specification, the decline in employment is particularly strong and statistically significant for later-born siblings in three-child families, with declines of 9 and 13 percentage points. Tables C.3 and C.4 show that, again, the effects are very different in families with and without evidence of perfect fertility control.[12] Also, once we control for pregnancy intention status as in Table C.5, the birth order gradient in employment is attenuated relative to that in Table C.2. Further, in line with our education findings, column (3) in Tables C.6 and C.7 show that the reported decline in employment is statistically significant in the pro-life sample but not in the pro-choice sample.

[12] In the pooled specification, the effects for the third child are statistically significantly different from each other across the two tables.

3.5 Conclusions

In this paper, we connect two disjoint literatures to shed more light on a novel mechanism that can give rise to birth order effects. The sources of birth order effects have puzzled economists and social scientists for decades. We document that, as one might expect, the prevalence of unwanted children increases with birth order, particularly when moving from second- to third-born children. We then replicate earlier findings related to birth order effects on completed education for the United States using data from PSID and a research design that exploits within-family variation. We go on to show that these birth order effects are reduced once we account for the differential pregnancy intention status of children born into different birth order. Moreover, we show that birth order effects no longer arise once we focus on a subsample of children from families who had no unwanted children or on families with religion background with less stringent attitudes towards the use of abortion. We conclude that the increasing prevalence of unwanted children at higher parities could be an important mechanism behind the well-documented birth order effects.

We show that our results are robust to alternative hypotheses that have been proposed in the literature on birth order effects. In particular, we show that our results hold in intact families and that they are not likely the result of endogenous last-born effects arising from fertility stopping after the birth of an unhealthy child. We also investigate adult employment and find similar patterns of attenuation of negative birth order effects once we factor in pregnancy intention.

It is possible that families might be averse to inequality in outcomes among their offspring, and they may try to equalize these outcomes as a result. Yet, families that avoid unwanted births and do not exceed their target desired fertility seem to be in better position to equalize outcomes among their offspring. While further research is necessary to directly test this mechanism, our findings are consistent with it.

Taken together, our results suggest novel avenues for future research on birth order effects. It would be interesting to see whether birth order effects are stronger in countries with more imperfect fertility control. Similarly, it might be interesting to explore whether, within countries, birth order effects are stronger during periods where the ability to avoid unwanted births is more limited. By the same token, one would expect birth order effects to be more pronounced, everything else equal, among groups that for cultural reasons are less prone to utilize various forms of fertility control.

Table 3.1. Summary Statistics

	Mean	Standard Deviation	Min	Max
3-Child Family	0.52	0.50	0	1
First-Born	0.46	0.50	0	1
Second-Born	0.39	0.49	0	1
Third-Born	0.15	0.36	0	1
Unwanted	0.16	0.36	0	1
Completed Years of Education	13.71	2.15	1	17
Employed in 2011	0.86	0.34	0	1
Self Reported Good Health before 17	0.83	0.37	0	1
Male	0.51	0.50	0	1
Age	40.24	12.34	18	93
Mother's Age at Childbirth	24.78	5.23	13	48
Mother's Education in 1985	12.49	2.45	5	18
White	0.44	0.50	0	1
Black	0.19	0.39	0	1
Hispanic	0.02	0.14	0	1
Other Race	0.35	0.48	0	1
Observations	5499			

Notes: Sample includes children from families with 2 or 3 children, with non-missing values of the outcome variable. For some variables like mother's education and the indicator for pregnancy intention associated with each child, the effective sample size is smaller as observations with missing values are excluded from the summary statistics. Mother's education is collected as of 1985, the year in which the retrospective pregnancy history was collected.

Table 3.2. Patterns of Unwanted Fertility by Birth Order

	(1) 2-Child Families	(2) 3-Child Families
Birth Order = 1	0.11	0.13
	[0.09,0.12]	[0.11,0.15]
Birth Order = 2	0.15	0.16
	[0.13,0.17]	[0.14,0.19]
Birth Order = 3		0.37
		[0.33,0.40]
Observations	2361	2524

Notes: 95% confidence intervals in brackets. Sample: PSID children from families with two or three children whose maternal pregnancy intention status at conception is not missing.

Table 3.3. Percentage of Unwanted Children by Birth Order PSID and NSFG

	1985	2002	2006-10
Birth Order = 1	11.8	8.5	8.8
Birth Order = 2	15.6	11.3	11.3
Birth Order $\geq$ 3	36.6	26.6	23.0
Source	PSID	NSFG	NSFG

Notes: The table reports the percentage of children who are retrospectively assessed as unwanted by their mothers at the time of interview. A child is defined as unwanted if the mother reports that when the child was conceived, she was not planning to have any (more) children. Neither at that time, nor in the future. The retrospective information collected in 1985 is based on the 1985 PSID wave and its Childbirth and Adoption History File. The 1985 sample is limited to families with two or three children. It does not include one-child families, or any birth order higher than three. The information based on retrospective information collected in 2002 and 2006-2010 comes from a Child Trends (2013) report based on corresponding waves from the National Survey of Family Growth.

Table 3.4. Birth Order and Education - OLS

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families	(4) 2- and 3-Child Families	(5) 2-Child Families	(6) 3-Child Families
Birth Order = 2	-0.09*	-0.13*	-0.03	-0.31***	-0.33***	-0.29***
	(0.05)	(0.07)	(0.08)	(0.06)	(0.08)	(0.08)
Birth Order = 3	-0.25***		0.02	-0.44***		-0.45***
	(0.08)		(0.09)	(0.10)		(0.12)
Age				0.05***	0.04	0.06**
				(0.02)	(0.02)	(0.02)
Age Squared				-0.00**	-0.00	-0.00*
				(0.00)	(0.00)	(0.00)
Male				-0.46***	-0.51***	-0.41***
				(0.05)	(0.07)	(0.07)
Black				-0.42***	-0.41***	-0.41***
				(0.08)	(0.11)	(0.12)
Hispanic				-0.16	-0.29	-0.09
				(0.25)	(0.41)	(0.31)
Other Race				0.11	-0.08	0.22
				(0.37)	(0.49)	(0.55)
Family of 3 Children				-0.20***		
				(0.06)		
Constant	13.78***	13.98***	13.51***	7.99***	8.53***	7.19***
	(0.04)	(0.06)	(0.07)	(1.00)	(1.37)	(1.41)
Observations	5499	2652	2847	5499	2652	2847
Mean Dependent Variable	13.71	13.92	13.50	13.71	13.92	13.50

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates in columns (4)–(6) include indicators for various levels of maternal education, as well as and indicator for whether maternal education information is missing, maternal age at childbirth and dummy variables indicating various age categories. Dependent variable measures completed years of education.

Table 3.5. Birth Order and Education - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.24***	-0.16	-0.30***
	(0.08)	(0.13)	(0.10)
Birth Order = 3	-0.39***		-0.50***
	(0.15)		(0.18)
Male	-0.52***	-0.72***	-0.40***
	(0.06)	(0.10)	(0.08)
Constant	12.42***	14.00***	11.55***
	(1.29)	(1.69)	(1.77)
Observations	5499	2652	2847
Mean Dependent Variable	13.71	13.92	13.50

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. Dependent variable measures completed years of education.

Table 3.6. Birth Order and Education in Families with and without Perfect Fertility Control - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Panel A: Families *with* Evidence of Perfect Fertility Control			
Birth Order = 2	0.06	0.11	-0.03
	(0.14)	(0.20)	(0.19)
Birth Order = 3	-0.02		-0.10
	(0.25)		(0.32)
Observations	1904	1139	765
Mean Dependent Variable	14.25	14.39	14.04
Panel B: Families *without* Evidence of Perfect Fertility Control			
Birth Order = 2	-0.46***	-0.46***	-0.44***
	(0.10)	(0.16)	(0.13)
Birth Order = 3	-0.64***		-0.72***
	(0.18)		(0.22)
Observations	3595	1513	2082
Mean Dependent Variable	13.42	13.57	13.31
Panel C: Panel B minus *A* Differences [p-value]			
Birth Order = 2	-0.52***	-0.57**	-0.40*
	[0.002]	[0.027]	[0.078]
Birth Order = 3	-0.63**		-0.61
	[0.044]		[0.114]

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. Dependent variable measures completed years of education.

Table 3.7. Birth Order and the Probability of Being Unwanted - OLS

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families	(4) 2- and 3-Child Families	(5) 2-Child Families	(6) 3-Child Families
Birth Order = 2	0.04***	0.04***	0.03**	0.07***	0.08***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Birth Order = 3	0.25***		0.23***	0.28***		0.28***
	(0.02)		(0.02)	(0.02)		(0.02)
Age				0.00	-0.01**	0.01***
				(0.00)	(0.00)	(0.00)
Age Squared				0.00	0.00**	-0.00***
				(0.00)	(0.00)	(0.00)
Male				-0.02**	-0.02*	-0.03*
				(0.01)	(0.01)	(0.02)
Black				0.20***	0.19***	0.21***
				(0.02)	(0.02)	(0.02)
Hispanic				0.10***	0.04	0.14***
				(0.03)	(0.04)	(0.05)
Other Race				0.02	0.00	0.03
				(0.05)	(0.06)	(0.08)
Family of 3 Children				-0.02		
				(0.01)		
Constant	0.12***	0.11***	0.13***	0.68***	1.01***	0.29
	(0.01)	(0.01)	(0.01)	(0.22)	(0.28)	(0.31)
Observations	4885	2361	2524	4885	2361	2524
Mean Dependent Variable	0.17	0.13	0.21	0.17	0.13	0.21

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include an indicator for individuals whose mother's education information in 1985 is missing, maternal age at childbirth and dummy variables indicating various age categories. Dependent variable equals 1 if an individual is the result of an unwanted pregnancy, zero otherwise.

Table 3.8. Birth Order and the Probability of Being Unwanted - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	0.06***	0.04*	0.06***
	(0.02)	(0.02)	(0.02)
Birth Order = 3	0.24***		0.27***
	(0.03)		(0.04)
Male	-0.00	-0.02	0.01
	(0.01)	(0.02)	(0.02)
Constant	0.84***	0.68**	1.01***
	(0.24)	(0.31)	(0.36)
Observations	4885	2361	2524
Mean Dependent Variable	0.17	0.13	0.21

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. Dependent variable equals 1 if individual is the result of an unwanted pregnancy, zero otherwise.

Table 3.9. Birth Order and Education Accounting for Unwantedness - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.14	-0.01	-0.25**
	(0.09)	(0.14)	(0.12)
Birth Order = 3	-0.23		-0.37*
	(0.16)		(0.19)
Unwanted	0.05	-0.01	0.06
	(0.13)	(0.22)	(0.16)
Male	-0.53***	-0.71***	-0.41***
	(0.06)	(0.10)	(0.08)
Constant	12.33***	13.96***	11.45***
	(1.30)	(1.71)	(1.79)
Observations	5499	2652	2847
Mean Dependent Variable	13.71	13.92	13.50

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. For children with missing information on their maternal pregnancy status at conception we include an indicator which equals to one whenever the pregnancy intention information is missing and equals zero otherwise. Further we interact this indicator with the birth order indicators.

Table 3.10. Birth Order and Education - Family Fixed Effects (Pro-Life)

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.26**	-0.15	-0.32**
	(0.10)	(0.17)	(0.13)
Birth Order = 3	-0.40**		-0.51**
	(0.19)		(0.23)
Male	-0.38***	-0.62***	-0.24**
	(0.08)	(0.13)	(0.10)
Constant	13.57***	15.18***	12.81***
	(1.62)	(2.20)	(2.15)
Observations	3556	1673	1883
Mean Dependent Variable	13.62	13.80	13.47

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. Dependent variable measures completed years of education. Subsample of individuals from families with maternal religion *more* strongly associated with a pro-life stance on abortion.

Table 3.11. Birth Order and Education - Family Fixed Effects (Pro-Choice)

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.11	0.03	-0.20
	(0.14)	(0.23)	(0.18)
Birth Order = 3	-0.21		-0.33
	(0.26)		(0.30)
Male	-0.71***	-0.90***	-0.61***
	(0.11)	(0.18)	(0.14)
Constant	10.64***	11.81***	9.79***
	(2.42)	(3.10)	(3.01)
Observations	1943	979	964
Mean Dependent Variable	13.85	14.13	13.57

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. Dependent variable measures completed years of education. Subsample of individuals from families with maternal religion *less* strongly associated with a pro-life stance on abortion.

Table 3.12. Birth Order and Education within Intact Families - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.31**	-0.23	-0.38**
	(0.13)	(0.19)	(0.18)
Birth Order = 3	-0.62***		-0.69**
	(0.23)		(0.29)
Male	-0.57***	-0.80***	-0.39***
	(0.10)	(0.15)	(0.14)
Constant	10.67***	11.98***	10.21***
	(1.58)	(2.20)	(2.28)
Observations	1933	1004	929
Mean Dependent Variable	14.49	14.68	14.30

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. Dependent variable is completed years of education. Subsample of individuals from "intact" families as defined in Section 3.4.4.

Table 3.13. Birth Order and Good Health Before Age 17 - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.02	-0.01	-0.04
	(0.02)	(0.04)	(0.03)
Birth Order = 3	-0.05		-0.06
	(0.04)		(0.06)
Male	0.03	0.02	0.03
	(0.02)	(0.03)	(0.03)
Constant	1.26***	1.64***	1.04**
	(0.33)	(0.44)	(0.46)
Observations	3763	1868	1895
Mean Dependent Variable	0.82	0.83	0.82

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates include maternal age at childbirth and dummy variables indicating various age categories. Dependent variable is a binary indicator for a retrospective self-assessment of own health status earlier in life (before age 17). “Very Good” and “Excellent” health equal 1, zero otherwise.

APPENDICES

APPENDIX A

Appendix to How Family Size in Childhood Affects Long-Run Human Capital and Economic Opportunity

A.1 Theories on the Interaction between Child Quantity and Quality

Theoretically, a large literature in demography and economics examines the effect of family size on children's human capital. The quantity-quality trade-off model (Becker, 1960; Becker and Lewis, 1973; Willis, 1973) of fertility decisions and investment in children provides a framework for the study of family size. It conceptualizes parents' economic behavior when faced with a trade-off between having more children and investing more parental time and financial resources in each child. However, to date, empirical evidence on the effects of family size remains inconclusive (Duflo, 1998; Black, Devereux and Salvanes, 2005; Cáceres-Delpiano, 2006). Moreover, it is unclear whether the result from a particular IV strategy can be generalized to a broader group of individuals (Imbens and Angrist, 1994; Angrist, Lavy and Schlosser, 2010). The concern is that public policy cannot alter twinning or the gender of children, and the families that can be effectively targeted by policies may experience completely different effects than where the identification comes from. This concern seems to be confirmed by the inconsistent estimates across instruments. For instance, Black, Devereux and Salvanes (2010) find no significant effect of family size on children's IQ score when using sex composition as an instrument, but discover negative and significant estimates using twins instead.

In the classic quantity-quality model, all children are the results of lifetime fertility decisions made in a static setting. There are several proposed ways of expanding the quantity-quality model. One crucial addition would be to incorporate sequentiality in fertility decisions (Heckman and Willis, 1976) and the possibility of failure to avoid unwanted births after deciding the ex-ante optimal number of children (Michael and Willis, 1976). Empirical evidence on this front has been sparse, with the exception of Lin et al. (2017) who examine the distinction between planned and unplanned increase in family size.

Families that have imperfect ability to control fertility and end up having unwanted children after they decide to stop might experience substantial impacts on the children they already have. Unwanted births indicate disruption in parental plans regarding child investment, and has many documented negative impacts. A growing literature in economics and demography shows the negative life-cycle consequences of unwanted children. Many empirical studies leverage unique data sources that directly document pregnancy intention information from women (Rosenzweig and Wolpin, 1993; Joyce, Kaestner and Korenman, 2000, 2002; Miller, 2009; Lin and Pantano, 2015; Lin et al., 2017).

A.2 PSID Data and Sample Construction

My primary source of survey data is the Panel Study of Income Dynamics (PSID), a longitudinal survey of a nationally representative sample of U.S. individuals and families. The PSID cross-year individual file allows this paper to construct completed years of education, current and retrospective self-reported health status, and labor market participation and income in adulthood. Critical to this project, the PSID also contains valuable information about Head Start participation in childhood, complete childbirths and adoption history records by individual's parents, and geo-coded county of birth and county grew up reported by individuals.[1] Crucially, the complete childbirths history records allow me to link an individual to her mother, and to the mother's total number of children after age 40 and thus construct the individual's eventual total number of siblings after her mother has completed fertility.

The PSID data is merged with the source of variation in family size – access to legal abortion overtime across states and counties in the U.S. On the state level, in 1970, five states became the earliest in modern U.S. history where abortion was broadly available. Abortion was then legalized nationwide in 1973 after the Roe v. Wade decision. On the county level, I use the abortion service provider information provided by the Guttmacher Institute. The measure of county-level intensity accounts for within-state changes in the provision of abortion service between 1970 and 1979.

A.3 State-Level Variation in Abortion and Family Size

Appendix Table A.1 presents the magnitude of the estimates. Cohorts that are affected during the variation of abortion access are put into bins of being conceived between 1967 and 1969, and between 1970 and 1972.[2] Among all individuals and using my preferred specification (column 3 and column 6), having access to legal abortion one to three years after birth reduces number of younger siblings on average by 0.271. It also reduces chances

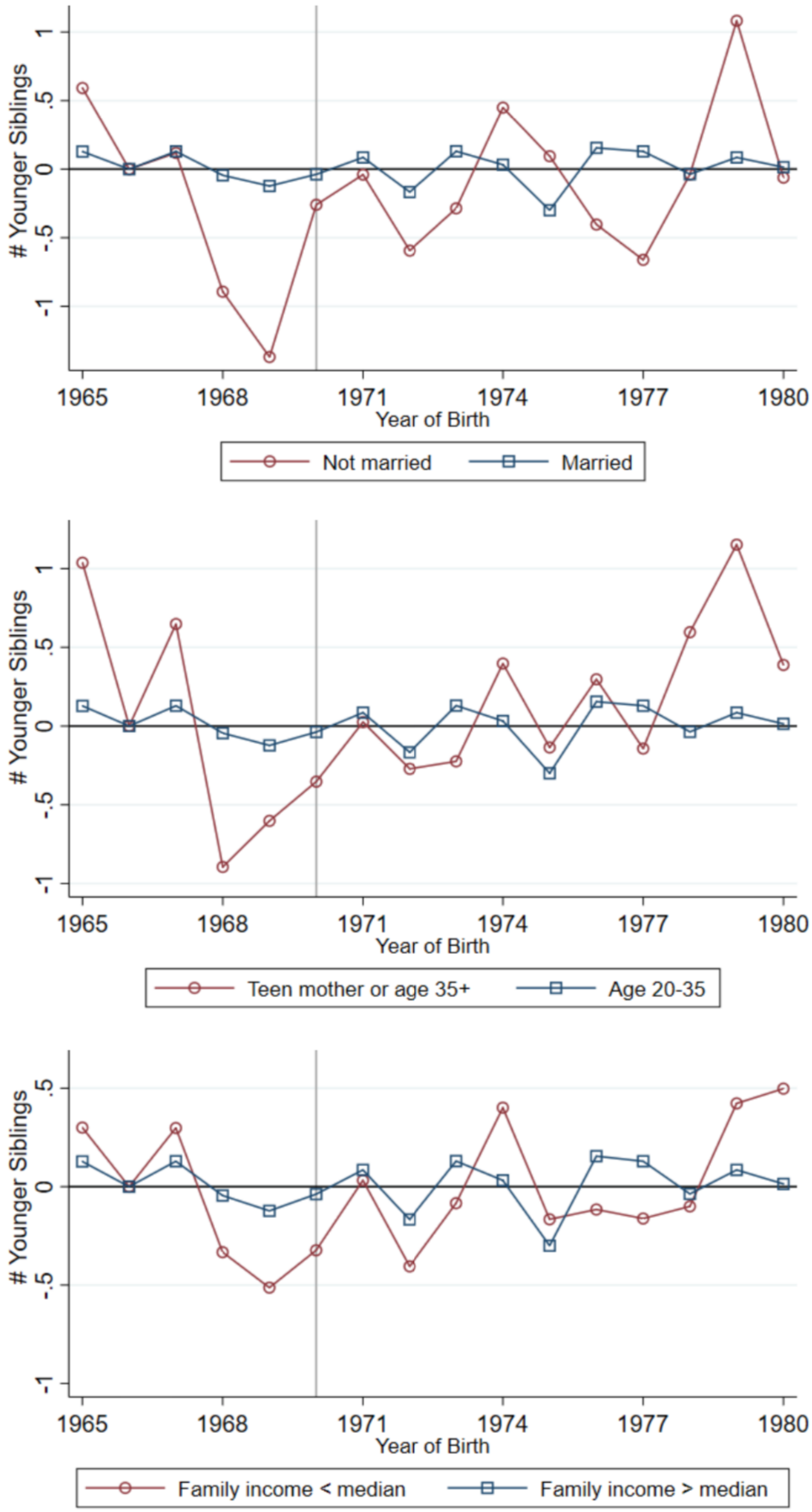
[1] County-level geo-code analysis requires approval to use PSID restricted data in the MICDA enclave.

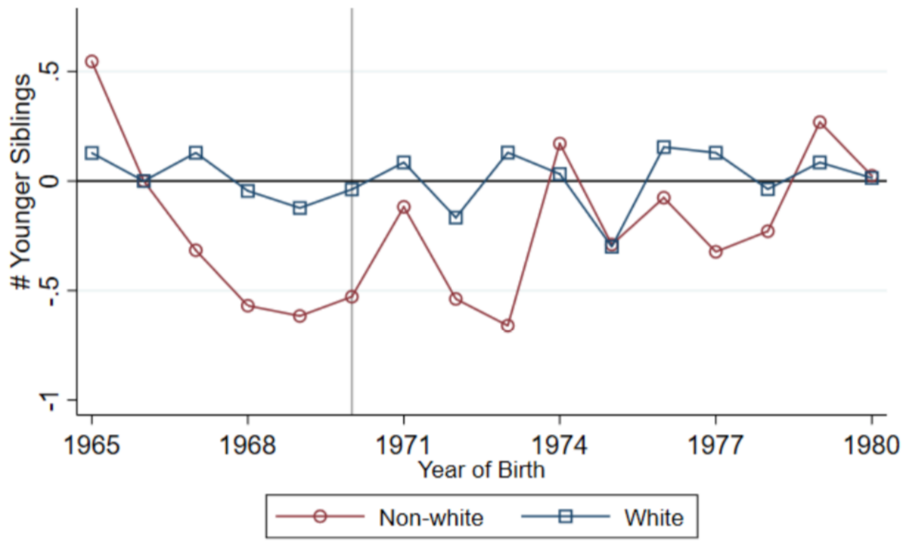
[2] Year of conception is constructed using year of birth and month of birth. It is set as one year before birth when the month of birth is between January and May. Although many children born after May are also conceived in the previous year, I assume for them access abortion is still possible if abortion becomes legal in their year of birth. The results are robust to different definitions of year conceived.

of ever having any younger sibling by 9 percentage points. The estimate appears to be robust to the addition of individual-level covariates and state and year of birth fixed effects. When stratifying by birth order, results appear to be similar albeit less precise, indicating a statistically significant treatment effect not driven by a change of composition of the families.[3]

[3] Note that the estimates tend to be smaller for the later-borns. This is not surprising given that they have fewer younger siblings than the first-borns on average.

Figure A.1. Difference in Number of Siblings, Early-Legalizing vs. Other States, by Family Background





Notes: Sample includes individuals in PSID cross-year individual file and Childbirth and Adoption History File whose mother last reported number of children after the age of 40. Point estimates of year-of-birth indicator interacted with repeal-state indicator are plotted. Covariates include gender, indicators of mother's race, mother's age at birth, and categorical variables of maternal completed education. Regression controls for year of birth and state of birth fixed effects. Confidence intervals presented are on the 90% level. Heteroskedasticity-robust standard errors clustered on the state level are used.

Table A.1. Abortion Legalization and Younger Siblings - Regression

	DV: # Younger Siblings			DV: 1 = Any Younger Siblings		
	(1)	(2)	(3)	(4)	(5)	(6)
A. All						
Conceived 1967-69 x Repeal	-0.250***	-0.243***	-0.271***	-0.082***	-0.081***	-0.090***
	[0.085]	[0.084]	[0.077]	[0.021]	[0.022]	[0.020]
Conceived 1970-72 x Repeal	-0.057	-0.047	-0.066	0.038	0.035	0.023
	[0.085]	[0.087]	[0.077]	[0.029]	[0.030]	[0.028]
B. First Births						
Conceived 1967-69 x Repeal	-0.176	-0.192	-0.243**	-0.032	-0.027	-0.032
	[0.134]	[0.130]	[0.117]	[0.032]	[0.035]	[0.033]
Conceived 1970-72 x Repeal	-0.033	-0.048	-0.085	0.094***	0.093***	0.079**
	[0.085]	[0.089]	[0.083]	[0.031]	[0.034]	[0.033]
C. Non-First Births						
Conceived 1967-69 x Repeal	-0.169	-0.171*	-0.178*	-0.058	-0.054	-0.060
	[0.101]	[0.101]	[0.100]	[0.042]	[0.040]	[0.038]
Conceived 1970-72 x Repeal	-0.173	-0.107	-0.115	-0.058	-0.037	-0.042
	[0.118]	[0.125]	[0.109]	[0.054]	[0.055]	[0.050]
Additional covariates	MGSY	MGESY	MGE	MGSY	MGESY	MGE
Year of Birth FE	No	No	Yes	No	No	Yes
State of Birth FE	No	No	Yes	No	No	Yes
Observations	6,456	6,435	6,434	6,456	6,435	6,434
Mean DV	1.100	1.100	1.100	0.630	0.630	0.630

Notes: Sample includes individuals in PSID cross-year individual file and Childbirth and Adoption History File born between 1965 and 1983 whose mother last reported number of children after the age of 40. Regressions use heteroskedasticity-robust standard errors clustered on the state level. Column 2 includes control variables for mother's race and age at birth M, individual's gender G, and indicator of born in early-repeal states S, and indicators of different birth cohort bins Y.

Table A.2. State-Level Abortion Ban Repeal and Human Capital

	Completed Yrs of Edu		High School or More		Some College or More		College Completion	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conceived 1967-69 x Repeal	0.365**	0.321**	-0.051*	-0.054**	0.037	0.021	0.158***	0.155***
	[0.146]	[0.140]	[0.027]	[0.025]	[0.038]	[0.036]	[0.044]	[0.044]
Conceived 1970-72 x Repeal	0.598***	0.578***	0.021	0.020	0.178***	0.164***	0.070	0.074
	[0.117]	[0.116]	[0.018]	[0.018]	[0.035]	[0.036]	[0.046]	[0.048]
Black	-0.565***	-0.523***	-0.031*	-0.022	-0.084***	-0.084***	-0.138***	-0.130***
	[0.106]	[0.104]	[0.017]	[0.015]	[0.025]	[0.023]	[0.021]	[0.024]
Female	0.688***	0.685***	0.043***	0.043***	0.152***	0.152***	0.110***	0.109***
	[0.046]	[0.046]	[0.009]	[0.009]	[0.014]	[0.014]	[0.012]	[0.012]
Maternal Age at Birth	0.051***	0.052***	0.004***	0.004***	0.007***	0.007***	0.012***	0.012***
	[0.006]	[0.006]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Maternal Yrs of Edu	0.347***	0.336***	0.024***	0.024***	0.066***	0.063***	0.064***	0.062***
	[0.017]	[0.017]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]
Year of Birth FE	No	Yes	No	Yes	No	Yes	No	Yes
State of Birth FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,358	4,357	4,358	4,357	4,358	4,357	4,358	4,357
Mean Y	13.70	13.70	0.910	0.910	0.590	0.590	0.300	0.300

Notes: Sample includes individuals in PSID cross-year individual file and Childbirth and Adoption History File born between 1965 and 1983 whose mother last reported number of children after the age of 40. Regressions use heteroskedasticity-robust standard errors clustered on the state level.

Table A.3. Abortion Provider Roll-Out and 1970 County Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	First service provider in					
	1970	1973	1974-75	1976-88	Ever	Never
Number of Counties	118	373	191	215	897	2138
Total population, 1970	275,765	238,770	69,748	54,524	163,485	20,458
% population growth 1960-1970	19.86	16.14	12.7	12.67	15.07	1.26
% population growth from net migration 1960-1970	9.14	3.34	2.19	1.01	3.3	-7.53
% female, 1970	50.4	51.13	50.83	50.67	50.86	50.73
% urban, 1970	54.78	61.72	46.6	42.82	53.06	26.38
% nonwhite, 1970	4.25	12.01	7.94	10.51	9.76	10.09
% of population aged 0-4 years, 1970	8.15	8.42	8.27	8.44	8.36	8.11
% of population aged 65+ years, 1970	10.5	9.4	11.08	10.49	10.16	12.77
Median age (years), 1970	29.06	27.91	29.15	28.42	28.45	30.56
Birth rate per 1,000 population, 1968	16.53	17.68	16.93	17.21	17.26	16.29
Death rate per 1,000 population, 1968	9.8	9.41	10.19	9.93	9.75	11.35
Median years schooling completed, persons 25+ (1970)	12.15	11.57	11.39	11.19	11.52	10.64
% persons 25+ with less than 5 yrs schooling (1970)	3.82	6.02	6.06	7.39	6.07	8.04
% persons 25+ w/ 12 or more yrs schooling (1970)	57.18	50.79	48.93	47	50.33	42.2
% female civilian labor force participation	27.69	29.84	27.01	27.31	28.35	24.5
% of families with female head, 1970	8.97	10.36	9.08	9.25	9.64	8.67
Median family income, all families, 1969 ($)	9649.88	8992.86	8147.29	7965.29	8652.95	6945.88
Leading party = Democratic	0.25	0.24	0.22	0.23	0.23	0.21

Notes: Unweighted averages are reported. Column 5 includes counties that had first abortion service provider between 1970 and 1988. *Source:* 1970 County and City Data Book.

Table A.4. State-Level Analysis of Family Size and Head Start Interaction

	(1) Completed Yrs of Edu	(2) High School or More	(3) Some College or More	(4) College Completion
Head Start x Conc 1967-69 x Repeal	1.02166*	-0.02786	0.23232*	-0.01049
	0.612	0.19	0.138	0.127
	0.096	0.884	0.094	0.934
Head Start x Conc 1970-72 x Repeal	-0.19414	0.86688***	0.13422	-0.64916***
	0.883	0.211	0.208	0.138
	0.826	0	0.519	0
Head Start	0.43378*	0.0453	0.08727	0.05282
	0.259	0.061	0.081	0.036
	0.095	0.461	0.285	0.145
Conc 1967-69 x Repeal	-0.07806	0.07392	-0.13891	0.16477**
	0.528	0.167	0.112	0.066
	0.883	0.658	0.215	0.014
Conc 1970-72 x Repeal	-0.12578	-0.54666***	-0.27619	0.37573***
	0.645	0.167	0.193	0.114
	0.846	0.001	0.155	0.001
Observations	20,108	20,108	20,108	20,108
R-squared	0.335	0.287	0.32	0.299
Y-mean	12.39	0.81	0.312	0.0896

Notes: Sample includes individuals in PSID cross-year individual file and Childbirth and Adoption History File born between 1960 and 1980 whose mother last reported number of children after the age of 40. Covariates include gender, indicators of mother's race, mother's age at birth, and categorical variables of maternal completed education. Regressions also include for indicator of Head Start interacted separately with repeal, conceived in 1967-69, conceived in 1970-72, as well as year of birth and state of birth fixed effects, resulting in a fully interacted specification. Confidence intervals presented are on the 90% level. Heteroskedasticity-robust standard errors clustered on the state level are used.

APPENDIX B

Appendix to The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency

B.1 Census Data

The project's primary data source is the 2000 Census and 2001 to 2013 ACS combined with the SSA's Numident file, accessed through project 1284 in the University of Michigan Research Data Center (RDC). The advantage of these data is that they link a rich set of productivity outcomes for cohorts potentially benefitting from War on Poverty programs (those who are ages 25 to 54 in the Census/ACS) with information on their access to Head Start programs in childhood using Numident information on county of birth. The 2000 census long-form contains information on 16.7 percent of the U.S. population; the 2001 to 2013 ACSs contain information for around 14 percent of the U.S. population. The number of Numidentlinked, unique individuals in these combined data sources represent about one-quarter of the U.S. population. In addition, we use the 1970 restricted long-form Census that contains information on school enrollment for children. Unfortunately, these data cannot be linked to the Numident because they have not yet been PIK'd by the Census.

B.2 SSA's Numident: Data on County and Date of Birth

Links from the Census/ACS to place and date of birth are important for studying the long-term impacts of Head Start, as place and date of birth provide crucial information on exposure of individuals to these programs in early childhood. For the Census/ACS files we use the survey-internal PIK code to match individuals with the SSA's Numident file. The Numident place-of-birth variable is a string variable detailing, in most cases, the city and state of birth. In previous work, Isen et al. (2013) developed a matching algorithm to connect this string variable to the Census Bureau database of places, counties, and minor civil divisions as well as the United States Geological Survey's Geographic Names Information System (GNIS) file. We also make use of code that was developed for a similar purpose by Black et al. (2015). Using both sources, we constructed a crosswalk between the NUMIDENT place of birth string variable and (standard) county FIPS codes, with over 90 percent of individuals matched to their counties of birth. Taylor et al. (2016)'s technical memorandum has been posted with the Census Bureau and contains this information.

The Census/ACS data have the benefit of providing a wide range of outcomes of interest, including individual earnings, but also program participation, disability, living arrangements, and family and household variables such as income and poverty. Additionally, we observe individuals in all states and we observe individuals regardless of whether they are employed. The Census/ACS data also have limitations. They are repeated cross-sections and information is self-reported (and, so, measured with error).

B.3 Data on School Age Entry Cutoffs

We restrict our sample to individuals who were born in areas where we have information for the relevant school-entry age cut-offs. This information is taken from Bedard and Dhuey

(2012) and supplemented using our own research for two states. According to state legislative documents, the school-entry age cutoff for the entire state of Texas was September 1st starting in 1969.[1] The state of Kansas' cutoff date is January 1st before 1965.[2]

We omit areas from our analysis sample where we are missing information on school entry cutoffs. This includes all individuals born in Colorado, Georgia, Indiana, Massachusetts, Montana before 1979, New Jersey, Rhode Island before 1967, South Carolina before 1978, Texas before 1969, Utah, Washington before 1977, and West Virginia before 1972. In our robustness checks, we find that the inclusion of these states and cohorts using a generic school-entry age cutoff tends to attenuate the estimates, as one might expect in the case of classical measurement error.

B.4 Data on Head Start

To study the long-run impacts of access to Head Start, we also use additional county-level sources of data to account for potentially confounding local programs and the economy. These data include information from the following sources: Bailey and Goodman-Bacon (2015) collected data on the OEO's community programs from the National Archives Community Action Program (NACAP) files, as well as from some administrative sources. For Community Health Centers, some information was hand-entered from annual Public Health Service (PHS) Reports. The resulting database contains information on (1) the county where a program delivered services, which allows each federal grant to be linked to county-level mortality rates; (2) the date that each county received its first program services grant (this excludes planning grants), which provides the year that programs began operating; and (3) some information on program grants between 1978 and 1980 from the National Archives

[1] http://www.heinonline.org/HOL/Page?men_tab=srchresults&handle=hein.ssl/sstx0163&size=2&collection=ssl&id=718

[2] http://www.heinonline.org/HOL/Page?men_tab=srchresults&handle=hein.ssl/ssks0074&size=2&collection=ssl&id=477

Federal Outlays (NAFO) files. We supplement these data with information on the legal services program from Cunningham (2013) and Food Stamps from Hoynes and Schanzenbach (2009).

B.5 Data on Head Start Launch Dates

Our main policy variable is the availability of the Head Start program, which was rolled out across counties during the War on Poverty. Bailey and Duquette (2014) and Bailey and Goodman-Bacon (2015) have compiled information from the National Archives and Records Administration on changes in Head Start funding between 1965 and 1980, which we use in this study. To verify their accuracy, these data have been compared to federal government directories of Head Start programs (Project Head Start 1971, Office of Child Development 1973, Project Head Start 1978). The following tables and figures supplement the analysis and description in the paper with more information on the roll-out of the program.

Table B.1. Share of Counties and Children under 6 in County with Head Start, 1965-1980

Fiscal Year HS Launched	# of Counties	Percent of counties	Cumulative percent of counties	Percent of kids age under 6 in 1970	Cumulative percent of kids age under 6 in 1970
1966	536	17.53%	17.57%	54.71%	55.22%
1967	217	7.10%	24.66%	10.57%	65.79%
1968	607	19.86%	44.52%	16.06%	81.85%
1969	41	1.34%	45.86%	0.71%	82.56%
1970	45	1.47%	47.33%	0.87%	83.43%
1971	10	0.33%	47.66%	0.17%	83.61%
1972	30	0.98%	48.64%	0.51%	84.12%
1973	7	0.23%	48.87%	0.11%	84.23%
1974	9	0.29%	49.17%	0.15%	84.38%
1975	16	0.52%	49.69%	0.29%	84.67%
1976	7	0.23%	49.92%	0.08%	84.75%
1977	4	0.13%	50.05%	0.16%	84.90%
1978	9	0.29%	50.34%	0.31%	85.21%
1979	3	0.10%	50.44%	0.08%	85.29%
No HS<1980	1515	49.56%	100.00%	14.71%	100.00%
Total	3057				

Table B.**2.** 1960 County Characteristics and Head Start's Launch, 1965-1980

	(1)	(2)	(3)	(4)	(5)
	Head Start launches before 1980	No Head Start before 1980	First Head Start grant in		
			1965-1966	1967-1968	1969-1980
# of Counties	1542	1515	537	824	181
County population	87,460	15,326	161,162	52,056	29,971
% Urban	72.21	31.02	79.44	61.4	42.35
% Rural Farm	5.72	24.44	3.42	9.01	16.38
% Nonwhite	10.77	11.2	11.33	9.45	12.31
% Population Aged 0-4	11.53	11.15	11.51	11.56	11.58
% Population aged 65+	8.95	10.89	8.74	9.29	9.66
Median Family Income (1959)	5712.77	4145.36	5984.64	5311.17	4550.95
Active MD per 1,000 Pop	0.69	0.02	0.96	0.23	0.04
AMR, All Ages	955.45	925.51	963.51	941.67	935.72
Infant Mortality Rate	25.6	27.28	25.36	25.77	28.01
Any Med Students (1969)	0.25	0.01	0.35	0.1	0
% Family Income <$3000	20.41	35.85	18.02	23.85	31.39
% Family Income $10,000+	15.15	7.44	16.73	12.74	9.04
% w/ 12+ Yrs Schooling (Age 25+)	41.81	34.18	43.04	40.01	36.37
% w/ <4 Yrs Schooling (Age 25+)	8.22	10.8	7.79	8.74	10.94
Gov'ts Exp ($000s) per 1,000 Pop (1957)	143.29	128.68	152.01	128.96	117.34
Labor Force Participation	0.38	0.36	0.38	0.37	0.36
% Labor Force Unemployed	5.32	4.8	5.38	5.21	5.24

Notes: All values are population-weighted means, with the exception of average county population in row 1. Characteristics are for 1960 unless otherwise specified. All variables are taken from the 1960 County and City Databooks (Haines et al. 2010) and 1990 Area Resource Files (US DHHS 1994) except the following. Medicare variables are for 1966, taken from the County-level Medicare File (US SSA 1969-1977; US HFA 1978-1980). Data on Medicare expenditures were shared by Almond, Hoynes and Schanzenbach (2011). We also use the 1959 to 1988 Vital Statistics Multiple-Cause of Death Files (US DHHS and NCHS 2007) to compute mortality rates.

Table B.**3.** Regression Analysis of 1960 County Characteristics and Head Start's Launch, 1965-1980

	(1)	(2)	(3)
% Nonwhite	0.008	0.006	0.005
	[0.006]	[0.007]	[0.007]
% Population Aged 0-4	0.014	0.053	0.059
	[0.060]	[0.063]	[0.061]
% Population Aged 65+	0.002	0.021	-0.001
	[0.033]	[0.035]	[0.035]
Median Family Income (1959)	0.000	0.000	0.001
	[0.000]	[0.000]	[0.000]
Active MD per 1000 Pop	-0.755***	-0.540***	0.395***
	[0.192]	[0.182]	[0.145]
AMR, All Ages	0.000	-0.000	-0.000
	[0.001]	[0.001]	[0.001]
AMR, 1960-65 Change	-0.000	0.000	0.000
	[0.001]	[0.001]	[0.001]
Infant Mortality Rate	0.000	-0.000	-0.003
	[0.005]	[0.006]	[0.006]
Any Med Students (1969)	-0.238	-0.439**	-0.401**
	[0.187]	[0.211]	[0.190]
% Family Income <$3,000	0.028	0.017	0.023
	[0.027]	[0.028]	[0.027]
% Family Income $10,000+	-0.090**	-0.090*	-0.064
	[0.042]	[0.048]	[0.046]
% w/ 12+ Yrs Schooling (Age 25+)	0.013	0.020	0.012
	[0.014]	[0.017]	[0.017]
% w/ <4 Yrs Schooling (Age 25+)	0.001	0.007	0.011
	[0.017]	[0.020]	[0.020]
Gov'ts Exp ($000s) per 1000 Pop (1957)	0.002	0.003	0.001
	[0.002]	[0.002]	[0.002]
Labor Force Participation	5.325***	5.417***	2.863
	[1.789]	[1.813]	[1.797]
% Labor Force Unemployed	-0.068**	-0.061**	-0.040
	[0.027]	[0.028]	[0.027]
% Labor Force Male	0.054***	0.050***	-0.006
	[0.017]	[0.018]	[0.019]
Constant	1,959***	1,960***	1,969***
	[3.305]	[3.621]	[3.608]
Observations	1,510	1,510	1,510
R-squared	0.065	0.113	0.157
Covariates		S,U	S,U,P

Notes: We estimate each regression by ordinary least squares. The dependent variable is year of Head Start's launch. Heteroskedasticity robust standard errors are beneath the point estimates in brackets. The regressions exclude counties that never received Head Start funding and are unweighted. Characteristics are for 1960 unless otherwise specified. Covariates without point estimate in the table include state fixed effects (S), urban categories (U, 0, $0 < u \leq 25$, $25 \leq u < 50$, $50 \leq u < 75$, $75 \leq u \leq 100$, where u is the share of a county's population living in an urban area), and log population (P). See also Table B.2 notes.

Table B.4. Age at Launch by Cohort and Year Head Start Launched

	Head Start launch year														
	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Birth cohort															
1950	15														
1951	14	15													
1952	13	14	15												
1953	12	13	14	15											
1954	11	12	13	14	15										
1955	10	11	12	13	14	15									
1956	9	10	11	12	13	14	15								
1957	8	9	10	11	12	13	14	15							
1958	7	8	9	10	11	12	13	14	15						
1959	6	7	8	9	10	11	12	13	14	15					
1960	5	6	7	8	9	10	11	12	13	14	15				
1961	4	5	6	7	8	9	10	11	12	13	14	15			
1962	3	4	5	6	7	8	9	10	11	12	13	14	15		
1963	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1964	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1965	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1966	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1967		-1	0	1	2	3	4	5	6	7	8	9	10	11	12
1968			-1	0	1	2	3	4	5	6	7	8	9	10	11
1969				-1	0	1	2	3	4	5	6	7	8	9	10
1970					-1	0	1	2	3	4	5	6	7	8	9
1971						-1	0	1	2	3	4	5	6	7	8
1972							-1	0	1	2	3	4	5	6	7
1973								-1	0	1	2	3	4	5	6
1974									-1	0	1	2	3	4	5
1975										-1	0	1	2	3	4
1976											-1	0	1	2	3
1977												-1	0	1	2
1978													-1	0	1
1979														-1	0
1980															-1

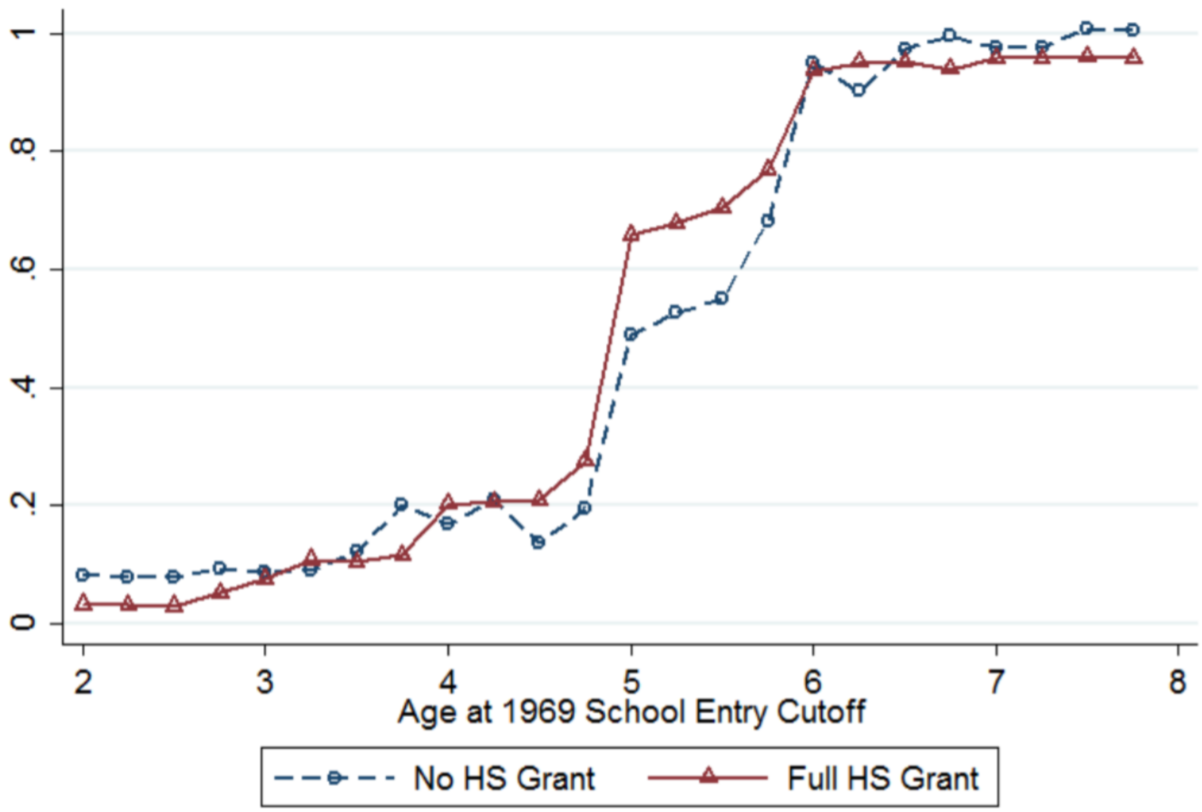
Notes: Table documents the age at Head Start's launch for each birth cohort and Head Start launch date in our data. Noteworthy is that our sample is compositionally balanced from ages -1 to 15, which we present in our event study graphical analysis. Outside of those ranges, the set of counties and birth cohorts will not be compositionally balanced.

B.6 The Effect of a Head Start Launch on Head Start Enrollment

Figure B.1 shows the unadjusted enrollment gap in the public IPUMS data between counties that had a Head Start program in 1970 versus those that did not. Notably, four-year-old children in counties without Head Start programs were 3.4 percentage points less likely to be enrolled in school (16.8pp versus 20.2pp). Five-year-old children were 17 percentage points less likely to be enrolled in school (48.9 versus 65.9). In the public data, these gaps are 5.9 percentage points among 4 year olds and 21.3 percentage points among 5 year olds when looking only at children of mothers with less than a high school education.[3]

[3]Note that Head Start was not exclusively for poor kids in the 1960s and 1970s. To encourage interaction between poor children and those from less disadvantaged backgrounds, OEO policy allowed 15 percent, and later 10 percent, of children to come from families that did not meet its poverty criteria. Roughly two-thirds of children in the full-year 1969 and 1970 programs came from families in which the mother had less than a high school education, although the mothers of about 7 percent of children had attended or graduated from college.

Figure B.1. 1970 School Enrollment by Mother's Education



Notes: The figure plots the predicted school enrollment by age for children in counties with Head Start versus those in counties without Head Start in 1970. These predictions come from a linear-probability model regression using a dependent variable is equal to one if a child was enrolled in school on February 1, 1970. Because the most detailed level of geography available in the public-use data is county group, availability of Head Start is operationalized as the population-weighted share of counties in an individual's county group that had Head Start by the 1969-1970 school year. The sample is limited to states where the school-entry cutoff falls at the beginning of a quarter and includes children between school age 2 and 7 using the school-entry age cutoff date in 1969. Source: Authors' calculations using the public 1970 Census (Ruggles et al., 2017), because we have not yet disclosed this figure in the restricted 1970 Census.

We explore these gaps in more detail using the restricted 1970 Census, which allows us to use a 1 in 6 sample of the U.S. population and county of residence (rather than county group). Using exact county rather than country group is more analogous to the long-term outcomes in the Census/ACS analysis. We compare school enrollment by a child's age in 1970 after adjusting for different county characteristics using either covariates or county fixed effects using the following specification:

$$SchoolEnrollment_{ic} = \boldsymbol{Z_c'}\boldsymbol{\beta_0} + \boldsymbol{A_i'}\boldsymbol{\beta_1} + \boldsymbol{A_i'}HeadStart_c\boldsymbol{\beta_2} + \epsilon_{ic} \tag{B.1}$$

where $\boldsymbol{A_i}$ has elements indicating the child's age relative to the school entry cut-off; $HeadStart_c$ is a dummy variable equal to 1 if the county had a Head Start program funded before 1970 (and is zero for places receiving their program in fiscal year 1971 or later). The set of covariates, $\boldsymbol{Z_c}$, includes either (1) $\theta_{s(c)}$, which captures state fixed effects to account for age-invariant, state-level factors that determine the local supply of preschools as well as 1960 county characteristics (share of county population in urban areas, in rural areas, under 5 years of age, 65 or older, nonwhite, with 12 or more years of education, with less than 4 years of education, in households with income less than $3,000, in households with incomes greater than $10,000, local government expenditures, income per capita, and whether the county was among the 300 poorest counties) or (2) county-level fixed effects (π_c). The point estimates of interest are the elements of $\boldsymbol{\beta_2}$, which, after regression-adjusting for county characteristics, capture differences in school enrollment rates of likely eligible children ages 4 to 5 in counties with Head Start.

Table B.5. Regression-Adjusted Relationship between Head Start and Enrollment

	(1)	(2)	(3)	(4)	(5)
	Dependent variable (DV): 1=School Enrollment				
	All Children			Boys	Girls
Head Start x age 3	0.038	0.038	0.038	0.037	0.039
	(0.0058)	(0.0058)	(0.0059)	(0.0065)	(0.0070)
Head Start x age 4	0.088	0.089	0.089	0.093	0.084
	(0.0088)	(0.0087)	(0.0088)	(0.0094)	(0.0095)
Head Start x age 5	0.148	0.149	0.149	0.151	0.145
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Regression-adjusted mean DV	0.52	0.56	0.52	0.52	0.52
Observations	830,000	830,000	830,000	420,000	410,000
State FE	Yes	Yes	No	No	No
County controls	No	Yes	No	No	No
County fixed effects	No	No	Yes	Yes	Yes

Notes: Sample is limited to children in states where we observe a school age entry cutoff, and where the school entry cutoff coincides with the beginning or end of a calendar quarter. Access to Head Start is measured as equal to 1 if a child lives in a county with a Head Start program in the 1969-70 school year. Observation counts are rounded per disclosure requirements. Omitted category is children age 7.75. Standard errors in brackets are clustered by county. Source: Authors calculations using the restricted 1970 Census.

The regression-adjusted, preschool enrollment gaps are summarized in Table B.5. School enrollment was 29 percent higher for all five year olds (0.149/0.52), 29 percent higher for boys (0.151/0.52) and 28 percent higher for girls (0.145/0.52). Although we are unable to assess many potential threats to the internal validity of this cross-sectional research design, the high degree of robustness of these estimates to the inclusion of different covariates in columns (2) and (3) is encouraging.

Our best estimate of the effect of a Head Start program from the 1970 Census on a birth cohort's exposure to Head Start is around 14.9 percentage points (Table B.5, column 3). It is 0.151 for men (column 4) and 0.145 for women (column 5). By construction, Census estimates should omit summer Head Start. An additional advantage of using the Census estimates is that they provide standard errors, which we use in our parametric bootstrap to scale our ITT estimates into ATETs.

Census estimates accord well with administrative data and are comparable to other studies. Administrative data suggest that in 1971, the average Head Start program served about 10 percent of resident 4-year-olds, which compares very well to around the 9 percent increase in school attendance in 1970 in the Census. The similarity of these administrative numbers and Census estimates suggests that crowd-out is minimal. To the extent that interested readers believe the estimate of the first stage should be higher or lower, they can deflate or inflate our ATET estimates accordingly.

B.7 Cost-Benefit Analysis of Head Start with the NLSY-79

A full accounting of the costs and benefits of Head Start is outside the scope of this paper. However, for comparison purposes, we compute the cumulative benefits of Head Start on economic opportunity through the program's cumulative effects on earnings potential. This is important in our context because Head Start appears to influence men and women's work effort, making the sample of wage earners selected. Our use of potential earnings follows Neal and Johnson (1996) and is directly comparable with Deming (2009). Like Deming, we use the National Longitudinal Survey of Youth 1979 (NLSY79) to predict wages for individuals born from 1957 to 1965 (ages 14 to 22 in 1979) – a time frame that overlaps our Census/ACS analysis. The NLSY data allow us to estimate the relationship between wages and components of the human capital and economic self-sufficiency indices after flexibly controlling for ability using the AFQT. Although AFQT is not available in the Census/ACS, using this as a covariate helps mitigate omitted variables bias in ability in the education and earnings relationship. We use observations on respondents' labor market wage income between 2002 and 2014, when they are between ages 35 and 57 years old, adjusted to be in 2013 dollars. We implement the following regression:

$$ln(Wages_i) = \boldsymbol{HC_i'}\boldsymbol{\beta_1} + \boldsymbol{ESS_i'}\boldsymbol{\beta_2} + \boldsymbol{X_i'}\boldsymbol{\beta_3} + \epsilon_i \tag{B.2}$$

where $\boldsymbol{X_i}$ is a vector of race, gender, age, survey year dummy variables and a quadratic in the respondent's standardized and age-normalized AFQT score (as in Deming (2009), Neal and Johnson (1996)). In some specifications, $\boldsymbol{X_i}$ includes Deming's covariates of age 19 outcomes as regressors to account for potential sources of omitted variables bias, which has a negligible effect on our calculations. We also include new outcomes that were contained in our human capital (HC) and economic self-sufficiency indices (ESS).

Note, however, that we omit components of the ESS that are directly related to log wages such as poverty and log family income. The resulting regression coefficients capture the importance of each index component after accounting flexibly for AFQT. Table B.6 (next page) shows that the results are generally very similar with and without Deming's covariates (columns 1-3 versus columns 4-6). Panel A summarizes the internal rate of return (IRR) using the estimated cost of Head Start per student of around \$5,400 and either the (1) realized human capital and self-sufficiency gains at ages 25 to 64 or the (2) savings in public assistance outlays. We present the regression estimates underlying the calculations in (1) in panel B.

In terms of the human capital and self-sufficiency returns only, we find an IRR to Head Start of 7.7 percent averaged over men and women (Panel A, column 1). The IRR ranges from around 4 percent for women to 11 percent for men (columns 2-3), owing both to the fact that women's human capital gains are smaller and that their labor-force effort falls in response to Head Start. Note, that this approach calculates only some of the private benefits that accrue to individuals and does not include benefits through improvements in outcomes not measured here. For instance, the extent to which more education engenders better health, longevity, or well-being is ignored in these calculations.

The human capital and self-sufficiency calculations also ignore savings on public assistance expenditures. The 2000 Census and 2001-2013 ACS suggest that the average amount of dollars received by public assistance recipients between ages 25 and 54 was $8,700 per year in 2013 dollars, which is 15 percent smaller than in the Survey of Income and Program Participation (SIPP) of $9,967 due to misreporting. Using the SIPP calculation, the IRR on putting one child through Head Start is 2.4% overall if the only returns to Head Start were in savings in public assistance expenditures: 2.5% for men and 2.2% for women.

Adding increases in tax revenues or subtracting deadweight loss encumbered by redistributing these expenditures through the tax and welfare system would serve to increase these estimates.

Table B.**6.** The Effect of Human Capital and Self-Sufficiency on Adult Earnings Potential

	(1)	(2)	(3)	(4)	(5)	(6)
		A. Ages 14 to 22			*B. Deming's sample: Ages 14 to 19*	
	Total	Male	Female	Total	Male	Female
A. Internal rate of return (IRR)						
HC and ESS only	7.7%	11.3%	4.0%	7.8%	11.6%	4.2%
Pub. assistance savings only	2.4%	2.5%	2.2%			
B. Regression estimates						
High School graduation	0.138***	0.0913*	0.217***	0.164***	0.116**	0.233***
	(0.035)	(0.048)	(0.053)	(0.041)	(0.055)	(0.065)
College completion	0.152***	0.215***	0.121***	0.158***	0.211***	0.135***
	(0.032)	(0.049)	(0.041)	(0.039)	(0.060)	(0.051)
Prof or doc degree	0.017	0.036	0.008	-0.0014	0.0228	0.003
	(0.036)	(0.060)	(0.044)	(0.043)	(0.070)	(0.054)
Professional job	0.249***	0.222***	0.256***	0.27***	0.225***	0.288***
	(0.016)	(0.025)	(0.020)	(0.020)	(0.028)	(0.024)
Years of schooling	0.030***	0.027**	0.027***	0.031***	0.0340***	0.0219*
	(0.007)	(0.011)	(0.009)	(0.009)	(0.013)	(0.012)
Weeks worked last year	0.035***	0.034***	0.036***	0.036***	0.0342***	0.0359***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Usual weekly hours	0.0199***	0.0156***	0.024***	0.020***	0.0153***	0.0240***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
AFQT score	0.190***	0.183***	0.202***	0.154***	0.153***	0.156***
	(0.011)	(0.015)	(0.015)	(0.012)	(0.017)	(0.018)
AFQT squared	-0.050***	-0.036***	-0.0655***	-0.037***	-0.036***	-0.0427***
	(0.007)	(0.010)	(0.011)	(0.008)	(0.011)	(0.012)
Idle				0.044***	0.045***	0.048***
				(0.009)	(0.014)	(0.012)
Crime				0.015	0.0179*	-0.002
				(0.009)	(0.011)	(0.020)
Teen pregnancy				-0.003	-0.005	0.004
				(0.008)	(0.015)	(0.010)
Health				0.011	0.011	0.011
				(0.009)	(0.014)	(0.011)
Constant	7.843***	7.872***	6.877***	7.818***	7.772***	6.912***
	(0.0947)	(0.139)	(0.120)	(0.112	(0.161	(0.147
Observations	36,536	18,252	18,284	25,508	13,004	12,504
# Individuals	7,327	3,653	3,674	5,053	2,578	2,475
R-squared	0.409	0.356	0.400	0.414	0.366	0.405

Notes: The dependent variable is log wage income. Control variables not reported in the table include race, gender, and birth and survey year fixed effects. High school completion, college completion, and professional or doctoral degree indicate completed years of education is greater or equal to 12, 16, or 18, respectively. Heteroskedasticity-robust standard errors are beneath the point estimates in parentheses. Standard errors are clustered at the individual level to account for longitudinal dependence in the data.

B.8 Additional Estimates

The following tables present estimates by race and sex for both human capital and economic self-sufficiency. We omit these from the paper because they are imprecise for nonwhites, largely owing to the fact that nonwhite children comprise around 15 percent of the sample. In addition, smaller sample sizes for these subgroups suggest relying on the parameterized spline estimate rather than the event-study point estimate at age -1.

Table B.7. The Effect of Head Start on Adult Human Capital by Race

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variables	Control mean (std.)	Event study at -1 (s.e.)	ITT % Gain	Spline at -1	F-stat on break at 6 [p-value]	ATET [95% CI]	ATET %Gain
A. Whites							
Human capital index	0.01	0.0132		0.014	11.61	.089	
	(0.203)	(0.0036)			[0.00]	[.04,.154]	
High school or GED completed	0.93	0.0355	3.8	0.0021	1.479	.238	26%
	(0.076)	(0.0125)			[0.22]	[.071,.449]	
Attended some college	0.63	0.0024	0.4	0.007	4.039	.016	2.5%
	(0.144)	(0.0011)			[0.04]	[.002,.034]	
Completed 4 year college	0.3	0.0076	2.5	0.0084	7.918	.051	17%
	(0.127)	(0.0023)			[0.00]	[.02,.091]	
Prof. or doc. Degree	0.029	0.0063	21.7	0.00175	6.957	.042	145%
	(0.039)	(0.0026)			[0.01]	[.008,.085]	
Years of schooling	13.64	0.002	0.0	0.0441	10.46	.013	0.10%
	(0.708)	(0.0006)			[0.00]	[.005,.024]	
Has a professional job	0.36	0.0067	1.9	0.00658	4.645	.045	12.5%
	(0.126)	(0.0024)			[0.03]	[.013,.085]	
B. Nonwhites							
Human capital index	0.015	0.0037		0.00147	5.278	.025	
	(0.33)	(0.0075)			[0.02]	[-.076,.13]	
High school or GED completed	0.86	-0.003	-0.3	-0.00004	2.134	-.02	-2.3%
	(0.16)	(0.0257)			[0.14]	[-.37,.34]	
Attended some college	0.56	-0.00018	0.0	0.000693	1.759	-.001	-0.2%
	(0.23)	(0.0035)			[0.18]	[-.049,.047]	
Completed 4 year college	0.19	0.00056	0.3	0.00126	5.126	.004	2.1%
	(0.18)	(0.0049)			[0.02]	[-.063,.073]	
Prof. or doc. Degree	0.017	0.004	23.5	0.00077	1.9	.027	159%
	(0.06)	(0.0044)			[0.17]	[-.031,.091]	
Years of schooling	13	0.0008	0.0	-0.007	2.712	.005	0.038%
	(1.16)	(0.0014)			[0.10]	[-.013,.025]	
Has a professional job	0.26	0.0028	1.1	0.000413	1.629	.019	7.3%
	(0.20)	(0.0049)			[0.20]	[-.047,.089]	

Notes: In column 1, the control mean and standard deviation are calculated using the cohorts ages 6 and 7 at the time Head Start was launched. Column 2 presents the estimated intention-to-treat (ITT) effect evaluated at birth cohort of full exposure (-1, see Figure 2.2). Columns 3 and 7 compute the percentage increase implied by the ITT or ATET, respectively, estimate relative to the control mean (the ratio of column 2 or 6 to column 1) for components of the index. Column 4 presents the ITT spline estimate evaluated at -1. Column 5 presents the F-statistic and p-value for the test of a trend-break in the spline at age 6. The ATET estimate in column 6 divides the ITT effect at -1 by the estimate of receiving a Head Start grant on school enrollment at school age 5 – 0.149 (s.e. 0.022) for the full sample, see Appendix Table B.5).

Table B.8. The Effect of Head Start on Adult Human Capital by Race: White Men and Women

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables	Control mean (std.)	Event study at - 1 (s.e.)	ITT % Gain	Spline at -1	F-stat on break at 6 [p-value]	ATET [95% CI]
A. White females						
Human capital index	0.011	0.0066		0.0098	3.57	.046
	(0.2328)	(0.0044)			[0.06]	[-.014,.115]
High school or GED completed	0.94	0.0017	0.2	0.00175	0.4132	.012
	(0.0827)	(0.0013)			[0.52]	[-.006,.032]
Attended some college	0.66	0.0038	0.6	0.00476	0.2301	.026
	(0.1668)	(0.0027)			[0.63]	[-.01,.069]
Completed 4 year college	0.31	0.0017	0.5	0.00546	1.8338	.012
	(0.1538)	(0.0032)			[0.18]	[-.032,.059]
Prof. or doc. Degree	0.023	0.0011	4.8	0.00084	3.6028	.008
	(0.0441)	(0.0007)			[0.06]	[-.002,.019]
Years of schooling	13.69	0.0114	0.1	0.0301	2.9967	.079
	(0.8099)	(0.0151)			[0.08]	[-.128,.306]
Has a professional job	0.38	0.0041	1.1	0.00483	3.1632	.028
	(0.1559)	(0.0031)			[0.08]	[-.014,.077]
B. White males						
Human capital index	0.0098	0.0198		0.0182	11.78	.131
	(0.2515)	(0.0043)			[0.00]	[.072,.213]
High school or GED completed	0.92	0.0033	0.4	0.00259	1.7827	.022
	(0.1016)	(0.0016)			[0.18]	[.001,.047]
Attended some college	0.6	0.0115	1.9	0.0098	6.1983	.076
	(0.1791)	(0.0030)			[0.01]	[.036,.13]
Completed 4 year college	0.3	0.0108	3.6	0.0112	9.0044	.072
	(0.1560)	(0.0029)			[0.00]	[.033,.124]
Prof. or doc. Degree	0.036	0.0029	8.1	0.0028	3.9562	.019
	(0.0557)	(0.0009)			[0.05]	[.007,.035]
Years of schooling	13.59	0.0594	0.4	0.0588	11.0703	.393
	(0.8919)	(0.0146)			[0.00]	[.196,.663]
Has a professional job	0.35	0.0095	2.7	0.0091	2.8856	.063
	(0.1584)	(0.0027)			[0.09]	[.027,.11]

Notes: See Table B.7 notes.

Table B.9. The Effect of Head Start on Adult Human Capital by Race: Nonwhite Men and Women

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables	Control mean (std.)	Event study at -1 (s.e.)	ITT % Gain	Spline at -1	F-stat on break at 6 [p-value]	ATET [95% CI]
C. Nonwhite females						
Human capital index	0.015	0.009		0.00364	7.11	.062
	(0.3668)	(0.0091)			[0.01]	[-.062,.201]
High school or GED completed	0.87	-0.000041	0.0	0.00091	3.6186	0
	(0.1783)	(0.0042)			[0.06]	[-.06,.06]
Attended some college	0.6	0.0052	0.9	0.00301	3.4384	.036
	(0.2627)	(0.0061)			[0.06]	[-.048,.128]
Completed 4 year college	0.2	0.0088	4.4	0.00434	6.7319	.061
	(0.2117)	(0.0054)			[0.01]	[-.012,.147]
Prof. or doc. Degree	0.016	-0.00031	-1.9	0.0000098	0.5994	-.002
	(0.0640)	(0.0017)			[0.44]	[-.026,.022]
Years of schooling	13.1	0.0198	0.2	-0.00329	3.1776	.137
	(1.2751)	(0.0306)			[0.07]	[-.284,.594]
Has a professional job	0.29	0.0071	2.4	0.00147	3.381	.049
	(0.2335)	(0.0069)			[0.07]	[-.045,.155]
C. Nonwhite males						
Human capital index	0.016	-0.0039		-0.000322	0.546	-.026
	(0.4032)	(0.0105)			[0.46]	[-.171,.116]
High school or GED completed	0.85	-0.0014	-0.2	-0.00077	0.1913	-.009
	(0.2037)	(0.0051)			[0.66]	[-.079,.06]
Attended some college	0.51	-0.0054	-1.1	-0.00189	0.0057	-.036
	(0.2839)	(0.0072)			[0.94]	[-.138,.061]
Completed 4 year college	0.18	-0.0014	-0.8	-0.00182	0.3928	-.009
	(0.2188)	(0.0059)			[0.53]	[-.09,.071]
Prof. or doc. Degree	0.018	0.0014	7.8	0.00147	1.1057	.009
	(0.0739)	(0.0021)			[0.29]	[-.018,.039]
Years of schooling	12.87	-0.0352	-0.3	-0.0077	0.9593	-.233
	(1.4293)	(0.0356)			[0.33]	[-.746,.241]
Has a professional job	0.22	-0.00059	-0.3	0.000294	0.0229	-.004
	(0.2324)	(0.0062)			[0.88]	[-.088,.081]

Notes: See Table B.7 notes.

Table B.10. The Effect of Head Start on Adult Self-Sufficiency by Race

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables	Control mean (std.)	Event study at -1 (s.e.)	ITT % Gain	Spline at -1	F-stat on break at 6 [p-value]	ATET [95% CI]
A. Whites						
Economic self-sufficiency index	0.021	0.0055		0.0041	5.328	.037
	(0.2)	(0.0024)			[0.02]	[.005,.076]
Worked last year	0.87	0.001	0.11	0.0010	0.4083	.007
	(0.1)	(0.0014)			[0.52]	[-.012,.027]
No. weeks worked last year	41.6	0.0558	0.13	0.0539	0.6303	.374
	(5.2)	(0.0758)			[0.43]	[-.636,1.47]
Usual hours works per week	36.2	0.0229	0.06	0.0455	1.275	.154
	(4.9)	(0.0804)			[0.26]	[-1.08,1.49]
In poverty*	0.08	-0.0015	-1.9	-0.0016	5.059	-.01
	(0.1)	(0.0011)			[0.02]	[-.027,.005]
Rec'd income from public sources*	0.1	-0.0052	-5.2	-0.0040	13.22	-.035
	(0.1)	(0.0010)			[0.00]	[-.056,-.02]
B. Nonwhites						
Economic self-sufficiency index	0.018	-0.0017		0.0070	1.188	-.011
	(0.28)	(0.0056)			[0.28]	[-.089,.066]
Worked last year	0.79	-0.0022	-0.28	0.0021	0.1255	-.015
	(0.20)	(0.0038)			[0.72]	[-.069,.037]
No. weeks worked last year	37.0	-0.069	-0.19	0.1792	1.211	-.463
	(10)	(0.2000)			[0.27]	[-3.255,2.29]
Usual hours works per week	32.4	-0.0551	-0.17	0.1099	0.2172	-.37
	(9.38)	(0.1820)			[0.64]	[-2.894,2.145]
Not in poverty	0.2	-0.0004	-0.20	-0.0020	0.9526	-.003
	(0.19)	(0.0038)			[0.33]	[-.055,.05]
No income from public sources	0.19	0.0013	0.68	-0.0003	0.0164	.009
	(0.19)	(0.0035)			[0.90]	[-.038,.058]

Notes: See Table B.7 notes.

Table B.11. The Effect of Head Start on Adult Self-Sufficiency by Race and Race-Sex: White Women and Men

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables	Control mean (std.)	Event study at -1 (s.e.)	ITT % Gain	Spline at -1	F-stat on break at 6 [p-value]	ATET [95% CI]
A. White females						
Economic self-sufficiency index	0.015	0.0063		0.0041	2.95	.043
	(0.1834)	(0.0030)			[0.09]	[.003,.093]
Worked last year	0.79	-0.0021	-0.3	-0.0022	0.2406	-.014
	(0.1346)	(0.0023)			[0.62]	[-.049,.018]
No. weeks worked last year	37.96	-0.123	-0.3	-0.1120	0.2377	-.848
	(6.9050)	(0.1160)			[0.63]	[-2.628,.761]
Usual hours works per week	30.56	-0.148	-0.5	-0.1197	0.1334	-1.021
	(5.9220)	(0.1040)			[0.72]	[-2.665,.414]
Log labor income wage only	10.31	0.0039		0.0048	2.33	.027
	(0.3501)	(0.0056)			[0.13]	[-.05,.111]
Log family income/poverty line	5.87	0.0069		0.0059	1.5198	.048
	(0.3097)	(0.0058)			[0.22]	[-.031,.137]
Not in poverty	0.9	0.0022	0.2	0.0025	5.966	.015
	(0.1026)	(0.0017)			[0.01]	[-.008,.042]
No income from public sources	0.89	0.0047	0.5	0.0025	0.7188	.032
	(0.1128)	(0.0015)			[0.40]	[.012,.059]
B. White males						
Economic self-sufficiency index	0.029	0.0045		0.0042	1.81	.030
	(0.2007)	(0.0031)			[0.18]	[-.01,.076]
Worked last year	0.92	0.004	0.4	0.0034	4.2597	.026
	(0.1011)	(0.0013)			[0.04]	[.009,.048]
No. weeks worked last year	45.5	0.195	0.4	0.1750	4.7986	1.291
	(5.7792)	(0.0694)			[0.03]	[.379,2.443]
Usual hours works per week	42.01	0.156	0.4	0.1694	5.54	1.033
	(5.8846)	(0.0818)			[0.02]	[-.025,2.292]
Log labor income wage only	10.92	0.0011		0.0014	0.0608	.007
	(0.3009)	(0.0046)			[0.81]	[-.054,.071]
Log family income/poverty line	5.95	0.00071		-0.0005	0.0218	.005
	(0.2961)	(0.0052)			[0.88]	[-.065,.077]
Not in poverty	0.94	0.00049	0.1	0.0002	0.0754	.003
	(0.0884)	(0.0013)			[0.78]	[-.014,.021]
No income from public sources	0.9	0.0053	0.6	0.0053	18.34	.035
	(0.1081)	(0.0014)			[0.00]	[.016,.06]

Notes: See Table B.7 notes.

Table B.12. The Effect of Head Start on Adult Self-Sufficiency by Race and Race-Sex: Nonwhite Women and Men

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables	Control mean (std.)	Event study at -1 (s.e.)	ITT % Gain	Spline at -1	F-stat on break at 6 [p-value]	ATET [95% CI]
C. Nonwhite females						
Economic self-sufficiency index	0.017	0.0016		0.0067	1.45	.011
	(0.3127)	(0.0070)			[0.23]	[-.087,.113]
Worked last year	0.75	-0.0019	-0.3	0.0032	0.8274	-.013
	(0.2278)	(0.0052)		)	[0.36]	[-.088,.061]
No. weeks worked last year	35.56	0.003	0.0	0.1855	1.5066	.021
	(11.8287)	(0.2690)			[0.22]	[-3.772,3.9]
Usual hours works per week	30.05	-0.0199	-0.1	0.0959	0.4792	-.137
	(10.1554)	(0.2280)			[0.49]	[-3.37,3.122]
Log labor income wage only	10.25	0.0014		0.0053	1.3066	.01
	(0.5443)	(0.0129)			[0.25]	[-.171,.196]
Log family income/poverty line	5.39	-0.0099		-0.0015	0.1804	-.068
	(0.5615)	(0.0141)			[0.67]	[-.278,.129]
Not in poverty	0.77	0.0016	0.2	0.0012	1.5023	.011
	(0.2220)	(0.0052)			[0.22]	[-.061,.087]
No income from public sources	0.8	-0.00008	0.0	0.0028	0.3556	-.001
	(0.2184)	(0.0044)			[0.55]	[-.063,.063]
D. Nonwhite males						
Economic self-sufficiency index	0.017	-0.0113		0.0049	0.055	-.075
	(0.3441)	(0.0083)			[0.82]	[-.198,.035]
Worked last year	0.82	-0.0066	-0.8	0.0012	0.0741	-.044
	(0.2209)	(0.0056)			[0.79]	[-.126,.031]
No. weeks worked last year	38.84	-0.334	-0.9	0.0490	0.0002	-2.212
	(12.1310)	(0.2920)			[0.99]	[-6.488,1.661]
Usual hours works per week	35.43	-0.274	-0.8	0.0105	0.1362	-1.815
	(11.4734)	(0.2750)			[0.71]	[-5.778,1.849]
Log labor income wage only	10.56	-0.00021		-0.0042	0.2214	-.001
	(0.5432)	(0.0134)			[0.64]	[-.182,.183]
Log family income/poverty line	5.57	0.0019		0.0036	0.1884	.013
	(0.5623)	(0.0124)			[0.66]	[-.154,.185]
Not in poverty	0.83	-0.003	-0.4	0.0016	0.0131	-.02
	(0.2156)	(0.0051)			[0.91]	[-.091,.049]
No income from public sources	0.82	-0.0026	-0.3	-0.0028	0.1825	-.017
	(0.2218)	(0.0053)			[0.67]	[-.091,.054]

Notes: See Table B.7 notes.

Table B.13. The Effect of Head Start on Incarceration and Mortality

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables	Control mean (std.)	Event study at -1 (s.e.)	ITT % Gain	Spline at -1	F-stat on break at 6 [p-value]	ATET [95% CI]
A. Full Sample						
Survived to 2000	0.98	-0.000089	-1.6	-0.0001	0.404	-.001
	(0.0057)	(0.0003)			[0.52]	[-.004,.003]
Incarcerated	0.01	0.00017	1.7	0.0002	0.111	.001
	(0.0277)	(0.0005)			[0.74]	[-.005,.008]
C. White females						
Survived to 2000	0.99	0.000016	0.4	0.0000	0.178	0
	(0.0044)	(0.0002)			[0.67]	[-.002,.003]
Incarcerated	0.00	0.00028	n/a	0.0003	2.62	.002
	(0.0140)	(0.0003)			[0.11]	[-.002,.006]
C. White males						
Survived to 2000	0.97	0.00022	2.7	0.0003	0.767	.001
	(0.0083)	(0.0003)			[0.38]	[-.003,.006]
Incarcerated	0.01	0.00019	1.9	0.0001	2.62	.001
	(0.0409)	(0.0008)			[0.59]	[-.009,.012]
D. Nonwhite females						
Survived to 2000	0.98	0.0003	2.7	-0.0001	3.43	.002
	(0.0111)	(0.0005)			[0.06]	[-.004,.009]
Incarcerated	0.01	0.00112	11.2	-0.0002	2.62	.008
	(0.0486)	(0.0017)			[0.62]	[-.016,.033]
E. Nonwhite males						
Survived to 2000	0.96	0.0000099	0.1	-0.0006	0.862	0
	(0.0176)	(0.0008)			[0.35]	[-.011,.011]
Incarcerated	0.08	0.000055	0.1	0.0014	2.62	0
	(0.1758)	(0.0062)			[0.88]	[-.083,.086]

Notes: See Table B.7 notes.

Table B.14. The Effect of Head Start on the Human Capital Index using Different Measures of Access

	(1) This paper (spline at -1)	(2) Barr and Gibbs (2017) differences-in-differences	(3) Years of access	(4) Thompson (2017) funding per capita	(5) Johnson and Jackson (2017) funding per capita
A. Excluding State-by-Birth-Year Fixed Effects					
Effect of Head Start	0.0152	0.0152	0.0191	0.00084	0.00130
	(0.00064)	(0.00477)	(0.00571)	(0.000458)	(0.000562)
B. Including State-by-Birth-Year Fixed Effect					
Effect of Head Start	0.0154	0.0110	0.0147	0.000347	0.000724
	(0.003)	(0.00256)	(0.00317)	(0.000439)	(0.000395)

Notes: There are 14,800,000 individuals (rounded for disclosure) in each regression. Panel A presents the results from specifications that exclude state-by-year-of birth fixed effects as is necessary in longitudinal samples. Panel B presents the results from specifications that exclude state-by-year-of-birth fixed effects as is necessary in longitudinal samples. The only change in the specification is the specification of the variable used to measure access to Head Start. Column 1 repeats the value shown in the paper. Column 2 presents a differences-in-differences specification, where Head Start=1 in a county for all children younger than 6 when it began. Column 3 uses a specification that measures share of the three potential years of access in which a child lived in a county with a Head Start program and was age eligible (possible variable values are 0, 1/3, 2/3, and 1). Column 4 uses the Thompson (2017) measure of Head Start access as average per-year, per-capita Head Start spending in the three years when an individual was between 3 and 6 years old. Column 5 uses the Johnson and Jackson (2017) measure of Head Start access as Head Start spending per poor 4-year-old, as defined when the cohort was 4 years old.

Table B.15. The Effect of Head Start on the Self-Sufficiency Index using Different Measures of Access

	(1) This paper (spline at -1)	(2) Barr and Gibbs (2017) differences-in-differences	(3) Years of access	(4) Thompson (2017) funding per capita	(5) Johnson and Jackson (2017) funding per capita
A. Excluding State-by-Birth-Year Fixed Effects					
Effect of Head Start	0.0055	0.00397	0.00488	0.000126	0.000193
	(0.00026)	(0.00146)	(0.00159)	(0.000201)	(0.000156)
B. Including State-by-Birth-Year Fixed Effect					
Effect of Head Start	0.0048	0.00232	0.00296	-0.000318	-1.85e-05
	(0.00022)	(0.00124)	(0.00147)	(0.000241)	(0.000119)

Notes: See Table B.14 notes.

Table B.16. The Effect of Head Start on Completed High School or GED using Different Measures of Access

	(1) This paper (spline at -1)	(2) Barr and Gibbs (2017) differences-in-differences	(3) Years of access	(4) Thompson (2017) funding per capita	(5) Johnson and Jackson (2017) funding per capita
A. Excluding State-by-Birth-Year Fixed Effects					
Effect of Head Start	0.0028	0.00357	0.00451	0.000164	0.000231
	(0.00026)	(0.00120)	(0.00140)	(0.000148)	(0.000108)
B. Including State-by-Birth-Year Fixed Effect					
Effect of Head Start	0.00245	0.00227	0.00305	-8.47e-05	6.73e-05
	(0.001)	(0.000673)	(0.000834)	(0.000144)	(7.52e-05)

Notes: See Table B.14 notes.

Table B.17. The Effect of Head Start on Enrolled in College using Different Measures of Accesss

	(1) This paper (spline at -1)	(2) Barr and Gibbs (2017) differences-in-differences	(3) Years of access	(4) Thompson (2017) funding per capita	(5) Johnson and Jackson (2017) funding per capita
A. Excluding State-by-Birth-Year Fixed Effects					
Effect of Head Start	0.0077	0.00949	0.0117	0.000309	0.000652
	(0.0006)	(0.00384)	(0.00467)	(0.000324)	(0.000299)
B. Including State-by-Birth-Year Fixed Effect					
Effect of Head Start	0.0077	0.00564	0.00733	5.23e-05	0.000211
	(0.002)	(0.00136)	(0.00175)	(0.000302)	(0.000170)

Notes: See Table B.14 notes.

APPENDIX C

Appendix to Birth Order and Unwanted Fertility

Table C.1. Birth Order and Employment - OLS

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families	(4) 2- and 3-Child Families	(5) 2-Child Families	(6) 3-Child Families
Birth Order = 2	-0.02	-0.01	-0.03	-0.03*	-0.02	-0.04*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Birth Order = 3	-0.04*		-0.04	-0.06**		-0.07**
	(0.02)		(0.03)	(0.03)		(0.03)
Age				0.03**	0.03	0.03*
				(0.01)	(0.02)	(0.02)
Age Squared				-0.00**	-0.00	-0.00
				(0.00)	(0.00)	(0.00)
Male				0.01	0.01	0.02
				(0.01)	(0.02)	(0.02)
Black				-0.12***	-0.12***	-0.11***
				(0.02)	(0.03)	(0.03)
Hispanic				-0.02	0.01	-0.04
				(0.05)	(0.06)	(0.07)
Other Race				0.05	0.11	-0.08
				(0.07)	(0.08)	(0.10)
Family of 3 Children				0.00		
				(0.02)		
Constant	0.87***	0.87***	0.87***	0.44	0.27	0.65
	(0.01)	(0.01)	(0.02)	(0.35)	(0.50)	(0.52)
Observations	2273	1169	1104	2273	1169	1104
Mean Dependent Variable	0.86	0.87	0.84	0.86	0.87	0.84

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Covariates in columns (4)-(6) include indicators for various levels of maternal education, as well as an indicator for whether maternal education information is missing. Age refers to age the child (now an adult) in the year 2011. Dependent Variable measures employment in 2011.

Table C.2. Birth Order and Employment - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.04	0.00	-0.09**
	(0.03)	(0.04)	(0.04)
Birth Order = 3	-0.05		-0.13**
	(0.05)		(0.06)
Male	0.01	-0.00	0.03
	(0.02)	(0.03)	(0.03)
Age	0.02	-0.02	0.05
	(0.02)	(0.04)	(0.03)
Age Squared	-0.00	0.00	-0.00*
	(0.00)	(0.00)	(0.00)
Constant	0.50	0.98	0.25
	(0.49)	(0.81)	(0.61)
Observations	2273	1169	1104
Mean Dependent Variable	0.86	0.87	0.84

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Age refers to age the child (now an adult) in the year 2011. Dependent Variable measures employment in 2011.

Table C.3. Birth Order and Employment in Families *with* Evidence of Perfect Fertility Control - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	0.01	0.07	-0.05
	(0.04)	(0.06)	(0.04)
Birth Order = 3	0.06		-0.08
	(0.08)		(0.08)
Male	0.02	0.01	0.05
	(0.03)	(0.04)	(0.05)
Age	0.04	0.06	0.01
	(0.04)	(0.06)	(0.07)
Age Squared	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)
Constant	-0.13	-0.92	0.97
	(0.89)	(1.25)	(1.32)
Observations	924	615	309
Mean Dependent Variable	0.90	0.91	0.89

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Age refers to age the child (now an adult) in the year 2011. Dependent Variable measures employment in 2011.

Table C.4. Birth Order and Employment in families *without* Evidence of Perfect Fertility Control - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.08*	-0.07	-0.11**
	(0.04)	(0.07)	(0.05)
Birth Order = 3	-0.13*		-0.15*
	(0.07)		(0.09)
Male	0.01	-0.03	0.02
	(0.03)	(0.06)	(0.04)
Age	0.03	-0.07	0.06
	(0.03)	(0.05)	(0.04)
Age Squared	-0.00	0.00	-0.00*
	(0.00)	(0.00)	(0.00)
Constant	0.59	2.32**	0.06
	(0.61)	(1.00)	(0.73)
Observations	1349	554	795
Mean Dependent Variable	0.83	0.82	0.83

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Age refers to age the child (now an adult) in the year 2011. Dependent Variable measures employment in 2011.

Table C.5. Birth Order and Employment Accounting for Unwantedness - Family Fixed Effects

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.04	-0.00	-0.07*
	(0.03)	(0.04)	(0.04)
Birth Order = 3	-0.02		-0.09
	(0.05)		(0.07)
Unwanted	0.06	0.13	0.04
	(0.05)	(0.09)	(0.07)
Male	0.01	0.00	0.02
	(0.02)	(0.03)	(0.03)
Age	-0.01	-0.01	0.01
	(0.03)	(0.04)	(0.04)
Age Squared	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)
Constant	0.99*	0.71	1.05
	(0.53)	(0.86)	(0.68)
Observations	2273	1169	1104
Mean Dependent Variable	0.86	0.87	0.84

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Age refers to age the child (now an adult) in the year 2011. Dependent Variable measures employment in 2011.

Table C.6. Birth Order and Employment - Family Fixed Effects (Pro-Life)

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.05	0.02	-0.11**
	(0.04)	(0.06)	(0.05)
Birth Order = 3	-0.06		-0.17**
	(0.07)		(0.08)
Male	-0.01	-0.03	0.01
	(0.03)	(0.05)	(0.04)
Age	0.02	-0.04	0.05
	(0.03)	(0.05)	(0.05)
Age Squared	-0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)
Constant	0.56	1.30	0.36
	(0.66)	(1.04)	(0.84)
Observations	1312	656	656
Mean Dependent Variable	0.86	0.86	0.87

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Age refers to age the child (now an adult) in the year 2011. Dependent Variable measures employment in 2011. Subsample of individuals from families with maternal religion *more* strongly associated with a pro-life stance on abortion.

Table C.7. Birth Order and Employment - Family Fixed Effects (Pro-Choice)

	(1) 2- and 3-Child Families	(2) 2-Child Families	(3) 3-Child Families
Birth Order = 2	-0.04	0.01	-0.08
	(0.05)	(0.07)	(0.06)
Birth Order = 3	-0.02		-0.08
	(0.08)		(0.12)
Male	0.02	0.03	0.01
	(0.03)	(0.04)	(0.05)
Age	0.03	0.04	0.04
	(0.05)	(0.07)	(0.06)
Age Squared	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)
Constant	0.23	0.02	0.25
	(0.98)	(1.51)	(1.29)
Observations	961	513	448
Mean Dependent Variable	0.84	0.87	0.81

Notes: Robust standard errors in parentheses, clustered at the family level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category is first-born child. Age refers to age the child (now an adult) in the year 2011. Dependent Variable measures employment in 2011. Subsample of individuals from families with maternal religion *less* strongly associated with a pro-life stance on abortion.

BIBLIOGRAPHY

BIBLIOGRAPHY

Almond, Douglas, and Janet Currie. 2011. "Human Capital Development before Age Five." Elsevier Handbook of Labor Economics.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the War on Poverty: The impact of Food Stamps on birth outcomes." *The Review of Economics and Statistics*, 93(2): 387–403.

Ananat, Elizabeth Oltmans, and Daniel M Hungerman. 2012. "The power of the pill for the next generation: oral contraception's effects on fertility, abortion, and maternal and child characteristics." *Review of Economics and Statistics*, 94(1): 37–51.

Ananat, Elizabeth Oltmans, Jonathan Gruber, and Phillip Levine. 2007. "Abortion legalization and life-cycle fertility." *Journal of Human Resources*, 42(2): 375–397.

Ananat, Elizabeth Oltmans, Jonathan Gruber, Phillip B Levine, and Douglas Staiger. 2009. "Abortion and selection." *The Review of Economics and Statistics*, 91(1): 124–136.

Angrist, Joshua D, and William N Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review*, 450–477.

Angrist, Joshua D, and William N Evans. 2000. "Schooling and labor market consequences of the 1970 state abortion reforms." In *Research in labor economics*. 75–113. Emerald Group Publishing Limited.

Angrist, Joshua, Victor Lavy, and Analia Schlosser. 2010. "Multiple experiments for the causal link between the quantity and quality of children." *Journal of Labor Economics*, 28(4): 773–824.

Arellano, M. 1987. "Computing Robust Standard Errors for Within-Groups Estimators." *Oxford Bulletin of Economics and Statistics*, 49(4): 431–434.

Argys, Laura M, Daniel I Rees, Susan L Averett, and Benjama (Kwan) Witoonchart. 2006. "Birth Order and Risky Adolescent Behavior." *Economic Enquiry*, 44(2): 215–233.

Averett, S L, L M Argys, and D I Rees. 2011. "Older siblings and adolescent risky behavior: does parenting play a role?" *Journal of Population Economics*, 24(3): 957–978.

Bagger, Jesper, Javier A. Birchenall, Hani Mansour, and Sergio Urzua. 2013. "Education, Birth Order, and Family Size." National Bureau of Economic Research Working Paper 19111.

Bailey, Martha, and Eric Chyn. 2020. "The Demographic Legacy of the Vietnam War: Evidence from the 1969 Draft Lottery." Working paper.

Bailey, Martha J. 2010. ""Momma's Got the Pill": How Anthony Comstock and Griswold v. Connecticut Shaped US Childbearing." *The American Economic Review*, 100(1): 98–129.

Bailey, Martha J. 2012. "Reexamining the Impact of Family Planning Programs on US Fertility: Evidence from the War on Poverty and the Early Years of Title X." *American Economic Journal: Applied Economics*, 4(2): 62–97.

Bailey, Martha J., and Andrew Goodman-Bacon. 2015. "The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans." *American Economic Review*, 105(3): 1067–1104.

Bailey, Martha J., and Nicolas J. Duquette. 2014. "How Johnson Fought the War on Poverty: The Economics and Politics of Funding at the Office of Economic Opportunity." *The journal of economic history*, 74(2): 351–388.

Bailey, Martha J., and Sheldon Danziger. 2013. "The Legacies of the War on Poverty." In *Legacies of the War on Poverty.* New York:Russell Sage Foundation.

Bailey, Martha J, Olga Malkova, and Zoë M McLaren. 2018. "Does Access to Family Planning Increase Childrens Opportunities? Evidence from the War on Poverty and the Early Years of Title X." *Journal of Human Resources*, 1216–8401R1.

Bailey, Martha, Shuqiao Sun, and Brenden Timpe. 2018. "'Prep School for Poor Kids': The Long-Run Impact of Head Start on Human Capital and Economic Self-sufficiency." Working paper.

Barmby, T, and A Cigno. 1990. "A sequential probability model of fertility patterns." *Journal of Population Economics*, 3(1): 3151.

Barnett, W.S., and Leonard N. Masse. 2007. "Comparative benefitcost analysis of the Abecedarian program and its policy implications." *Economics of Education Review*, 26(1): 113–125.

Barr, Andrew, and Chloe R Gibbs. 2017. "Breaking the Cycle? Intergenerational Effects of an Anti-Poverty Program in Early Childhood."

Bassok, Daphna, Maria Fitzpatrick, and Susanne Loeb. 2014. "Does state preschool crowd-out private provision? The impact of universal pre-kindergarten on the childcare sector in Oklahoma and Georgia." *Journal of Urban Economics*, 83: 18–33.

Baydar, Nazli. 1995. "Consequences for children of their birth planning status." *Family Planning Perspectives*, 27(6): 228–245.

Becker, Gary S. 1960. "An Economic Analysis of Fertility, Demographic and economic change in developed countries: a conference of the Universities." *National Bureau Commitee for Economic Research*, 209.

Becker, Gary S, and H Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2, Part 2): S279–S288.

Becker, Gary S, and Robert J Barro. 1988. "A reformulation of the economic theory of fertility." *The quarterly journal of economics*, 103(1): 1–25.

Behrman, Jere R. 1988. "Nutrition, health, birth order and seasonality: Intrahousehold allocation among children in rural India." *Journal of Development Economics*, 28(1): 43–62.

Behrman, Jere R., and Paul Taubman. 1986. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics*, 4(3, Part 2): S121–S145.

Behrman, Jere R, Robert A Pollak, and Paul Taubman. 1982. "Parental preferences and provision for progeny." *Journal of Political Economy*, 90(1): 52–73.

Belmont, Lillian, and Francis A. Marolla. 1973. "Birth Order, Family Size, and Intelligence." *Science*, 182(4117): 1096–1101.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Birdsall, Nancy. 1979. "Siblings and Schooling in Urban Colombia." PhD diss. Yale University.

Birdsall, Nancy. 1990. "Birth order effects and time allocation." *Research in Population Economics*, 7: 191–213.

Björkegren, Evelina, and Helena Svaleryd. 2017. "Birth Order and Child Health." Institute for Evaluation of Labour Market and Education Policy Working Paper.

Black, Sandra E, Erik Grönqvist, and Björn Öckert. 2017. "Born to lead? The effect of birth order on non-cognitive abilities." National Bureau of Economic Research Working Paper 23393.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *The Quarterly Journal of Economics*, 120(2): 669–700.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2010. "Small family, smart family? Family size and the IQ scores of young men." *Journal of Human Resources*, 45(1): 33–58.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2011. "Older and Wiser? Birth Order and IQ of Young Men." *CESifo Economic Studies*, 57(1): 103–120.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2016. "Healthy (?), wealthy, and wise: Birth order and adult health." *Economics & Human Biology*, 23: 27–45.

Bloom, Benjamin S. 1964. *Stability and change in human characteristics.* New York:Wiley. OCLC: 224354.

Booth, Alison L, and Hiau Joo Kee. 2009. "Birth order matters: the effect of family size and birth order on educational attainment." *Journal of Population Economics*, 22(2): 367–397.

Bound, John, and Gary Solon. 1999. "Double trouble: on the value of twins-based estimation of the return to schooling." *Economics of Education Review*, 18(2): 169–182.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement error in survey data." *Handbook of Labor Economics*, 5: 3705–3843.

Braun, Samuel J, and Esther P Edwards. 1972. *History and theory of early childhood education.* Worthington, Ohio, C.A. Jones Pub. Co. OCLC: 639993516.

Breining, Sanni, Joseph Doyle, David Figlio, Krzysztof Karbownik, and Jeffrey Roth. 2017. "Birth order and delinquency: Evidence from Denmark and Florida." National Bureau of Economic Research Working Paper 23394.

Brenøe, Anne A, and Ramona Molitor. 2018. "Birth order and health of newborns." *Journal of Population Economics*, 31(2): 397395.

Cáceres-Delpiano, Julio. 2006. "The impacts of family size on investment in child quality." *Journal of Human Resources*, 41(4): 738–754.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber. 2015. "Inference on Causal Effects in a Generalized Regression Kink Design." *Econometrica*, 83(6): 2453–2483.

Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy*, 6(4): 135–173.

Cascio, Elizabeth U., and Diane Whitmore Schanzenbach. 2013. "The Impacts of Expanding Access to High-Quality Preschool Education." *Brookings Papers on Economic Activity*, 127–178.

Charles, Kerwin Kofi, and Melvin Stephens, Jr. 2006. "Abortion legalization and Adolescent Substance use." *The Journal of Law and Economics*, 49(2): 481–505.

Child Trends. 2013. "Unintended Births."

Conley, Dalton, and Rebecca Glauber. 2006. "Parental Educational Investment and Childrens Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility." *Journal of Human Resources*, XLI(4): 722–737.

Cunha, Flavio, and James Heckman. 2007. "The technology of skill formation." *American Economic Review*, 97(2): 31–47.

Cunha, Flávio, Irma Elo, and Jennifer Culhane. 2013. "Eliciting maternal expectations about the technology of cognitive skill formation." National Bureau of Economic Research.

Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives*, 15(2): 213–238.

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review*, 85(3): 341–364.

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

De Haan, Monique. 2010. "Birth order, family size and educational attainment." *Economics of Education Review*, 29(4): 576–588.

De Haan, Monique, Erik Plug, and José Rosero. 2014. "Birth order and human capital development evidence from Ecuador." *Journal of Human Resources*, 49(2): 359–392.

Deming, David. 2009. "Early childhood intervention and life-cycle skill development: Evidence from Head Start." *American Economic Journal: Applied Economics*, 1(3): 111–134.

Dills, Angela K, and Jeffrey A Miron. 2006. "A Comment on Donohue and Levitts (2006) Reply to Foote and Goetz (2005)." *manuscript, Department of Economics, Harvard University.*

Donohue, John J, and Steven D Levitt. 2001. "The impact of legalized abortion on crime." *The Quarterly Journal of Economics*, 116(2): 379–420.

Donohue, John J, and Steven D Levitt. 2004. "Further evidence that legalized abortion lowered crime a reply to joyce." *Journal of Human Resources*, 39(1): 29–49.

Donohue, John J, and Steven D Levitt. 2008. "Measurement error, legalized abortion, and the decline in crime: A response to Foote and Goetz." *The Quarterly Journal of Economics*, 123(1): 425–440.

Donohue, John J, Jeffrey Grogger, and Steven D Levitt. 2009. "The impact of legalized abortion on teen childbearing." *American Law and Economics Review*, 11(1): 24–46.

Do, Quy-Toan, and Tung D Phung. 2010. "The importance of being wanted." *American Economic Journal: Applied Economics*, 2(4): 236–253.

Duflo, Esther. 1998. "Evaluating the effect of birth-spacing on child mortality." mimeo, Department of Economics, Massachusetts Institute of Technology.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Chapter 61 Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*. Vol. 4, , ed. T. Paul Schultz and John A. Strauss, 3895–3962. Elsevier.

Duncan, Greg J., and Katherine Magnuson. 2013. "Investing in Preschool Programs." *Journal of Economic Perspectives*, 27(2): 109–132.

Eckstein, Daniel, Kristen J Aycock, Mark A Sperber, John McDonald, V Van Wiesner III, Richard E Watts, and Phil Ginsburg. 2010. "A review of 200 birth-order studies: Lifestyle characteristics." *Journal of Individual Psychology*, 66(4): 408–434.

Edmonds, E V. 2006. "Understanding sibling differences in child labor." *Journal of Population Economics*, 19(4): 795–821.

Efron, Bradley, and Robert J Tibshirani. 1993. *An introduction to the bootstrap.* New York, N.Y.; London:Chapman & Hall. OCLC: 797437299.

Ejrnæs, Mette, and Claus C Pörtner. 2004. "Birth order and the intrahousehold allocation of time and education." *Review of Economics and Statistics*, 86(4): 1008–1019.

Ermisch, J F, and M Francesconi. 2001. "Family structure and children's achievements." *Journal of Population Economics*, 14(2): 249–270.

Espinosa, Linda M. 2002. "High-Quality Preschool: Why We Need It and What It Looks Like." 14.

Evans, John H. 2002. "Polarization in abortion attitudes in U.S. religious traditions, 19721998." *Sociological Forum*, 17(3): 397422.

Foote, Christopher L, and Christopher F Goetz. 2008. "The impact of legalized abortion on crime: Comment." *The Quarterly Journal of Economics*, 123(1): 407–423.

Fosburg, Linda B, Nancy N Goodrich, Mary Kay Fox, Patricia Granahan, Janet Smith, John H Rimes, and Michael Weitzman. 1984. "The Effects of Head Start Health Services: Report of the Head Start Health Evaluation." In *Report Prepared for the ADministration for Children, Youth and Families, U.S. Department of Health and Human Services*. 1191. Cambridge, MA:Abt Associates.

Foster, Andrew D, Mark R Rosenzweig, et al. 2007. "Does economic growth reduce fertility? Rural India 1971–99."

Frisvold, David E. 2015. "Nutrition and cognitive achievement: An evaluation of the School Breakfast Program." *Journal of Public Economics*, 124: 91–104.

Galton, Francis. 1875. *English men of science: Their nature and nurture.* D. Appleton.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *American Economic Review*, 92(4): 999–1012.

Gillezeau, Rob. 2010. "Did the War on Poverty cause race riots."

Goodman-Bacon, Andrew. 2018. "Public Insurance and Mortality: Evidence from Medicaid Implementation." *Journal of Political Economy*, 126(1): 216–262.

Griliches, Zvi. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy*, 87(5, Part 2): S37–S64.

Grosz, Michel Z., Douglas L. Miller, and Na'ama Shenhav. 2017. "All In the Family: Assessing the External Validity of Family Fixed Effects Estimates and the Long-Term Impact of Head Start." *Accessed July 31, 2017*.

Gruber, Jonathan, Phillip Levine, and Douglas Staiger. 1999. "Abortion legalization and child living circumstances: who is the marginal child?" *The Quarterly Journal of Economics*, 114(1): 263–291.

Hao, Lingxin, V Joseph Hotz, and Ginger Z Jin. 2008. "Games parents and adolescents play: Risky behaviour, parental reputation and strategic transfers." *The Economic Journal*, 118(528): 515–555.

Hechinger, Fred M., ed. 1966. *Pre-School Education Today: New Approaches to Teaching Three-, Four-, and Five-Year-Olds.* New York:Doubleday.

Heckman, James J, and Robert J Willis. 1976. "Estimation of a stochastic model of reproduction: An econometric approach." In *Household production and consumption.* 99–146. NBER.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The rate of return to the HighScope Perry Preschool Program." *Journal of Public Economics*, 94(1): 114–128.

Holm, Sture. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics*, 6(2): 65–70.

Horton, Susan. 1988. "Birth order and child nutritional status: evidence from the Philippines." *Economic Development and Cultural Change*, 36(2): 341–354.

Hotz, V. Joseph, and Juan Pantano. 2015. "Strategic Parenting, Birth Order and School Performance." *Journal of Population Economics*, 28(4): 911–936.

Hoxby, Caroline M. 2001. "All school finance equalizations are not created equal." *The Quarterly Journal of Economics*, 116(4): 1189–1231.

Hoynes, Hilary, Marianne Page, and Ann Huff Stevens. 2011. "Can targeted transfers improve birth outcomes?: Evidence from the introduction of the WIC program." *Journal of Public Economics*, 95(7): 813–827.

Hunt, Joseph McVicker. 1961. *Intelligence and Experience.* New York:Ronald Press.

Imbens, Guido W, and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467–475.

Johnson, Rucker C, and C. Kirabo Jackson. 2017. "Reducing Inequality Through Dynamic Complementarity: Evidence from Head Start and Public School Spending." National Bureau of Economic Research Working Paper 23489.

Joyce, Ted. 2004. "Did legalized abortion lower crime?" *Journal of Human Resources*, 39(1): 1–28.

Joyce, Ted, Ruoding Tan, and Yuxiu Zhang. 2013. "Abortion before & after Roe." *Journal of health economics*, 32(5): 804–815.

Joyce, Theodore J, Robert Kaestner, and Sanders Korenman. 2000. "The effect of pregnancy intention on child development." *Demography*, 37(1): 83–94.

Joyce, Theodore J, Robert Kaestner, and Sanders Korenman. 2002. "On the validity of retrospective assessments of pregnancy intention." *Demography*, 39(1): 199–213.

Kantarevic, Jasmin, and Stéphane Mechoulan. 2006. "Birth order, educational attainment, and earnings an investigation using the PSID." *Journal of Human Resources*, 41(4): 755–777.

Kaplan, John. 1988. "Abortion as a Vice Crime: A What If Story." *Law & Contemp. Probs.*, 51: 151.

Kessler, Daniel. 1991. "Birth order, family size, and achievement: Family structure and wage determination." *Journal of Labor Economics*, 9(4): 413–426.

Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *The Quarterly Journal of Economics*, 131(4): 1795–1848.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica*, 75(1): 83–119.

Kubička, L, Z Matějček, HP David, Z Dytrych, WB Miller, and Z Roth. 1995. "Children from unwanted pregnancies in Prague, Czech Republic revisited at age thirty." *Acta Psychiatrica Scandinavica*, 91(6): 361–369.

Lehmann, Jee-Yeon K., Ana Nuevo-Chiquero, and Marian Vidal-Fernandez. 2013. "Birth Order and Differences in Early Inputs and Outcomes." IZA Institute of Labor Economics Discussion Paper 6755.

Levine, Phillip B. 2004. *Sex and consequences: Abortion, public policy, and the economics of fertility.* Princeton University Press.

Levine, Phillip B, Douglas Staiger, Thomas J Kane, and David J Zimmerman. 1996. "Roe V. Wade and American Fertility." *NBER Working Paper*, , (w5615).

Levine, Robert A. 1970. *The Poor Ye Need Not Have With You: Lessons from the War on Poverty.* Cambridge:MIT Press.

Levitan, Sar A. 1969. *The Great Society's Poor Law: A New Approach to Poverty.* Baltimore:Johns Hopkins Press.

Lindert, Peter H. 1977. "Sibling position and achievement." *Journal of Human Resources*, 12(2): 198–219.

Lin, Wanchuan, and Juan Pantano. 2015. "The Unintended: Negative Outcomes over the Life Cycle." *Journal of Population Economics*, 28(2): 479–508.

Lin, Wanchuan, Juan Pantano, and Shuqiao Sun. 2019. "Birth order and unwanted fertility." *Journal of Population Economics.*

Lin, Wanchuan, Juan Pantano, R. Pinto, and S. Sun. 2017. "Identification of quantity quality tradeoff with imperfect fertility control." Unpublished working paper.

Loeb, Susanna. 2016. "Missing the target: We need to focus on informal care rather than preschool."

Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *The Quarterly Journal of Economics*, 122(1): 159–208.

Malamud, Ofer, Cristian Pop-Eleches, and Miguel Urquiola. 2016. "Interactions between family and school environments: Evidence on dynamic complementarities?" National Bureau of Economic Research.

Mechoulan, Stéphane, and François-Charles Wolff. 2015. "Intra-household allocation of family resources and birth order: evidence from France using siblings data." *Journal of Population Economics*, 28(4): 937–964.

Meyer, Bruce D., Wallace K.C. Mok, and James X. Sullivan. 2015. "Household surveys in crisis." *Journal of Economic Perspectives*, 29(4): 199–226.

Michael, Robert T, and Robert J Willis. 1976. "Contraception and fertility: Household production under uncertainty." In *Household Production and Consumption.* 25–98. NBER.

Miller, Amalia R. 2009. "Motherhood Delay and the Human Capital of the Next Generation." *American Economic Review*, 99(2): 154–58.

Mogstad, Magne, and Matthew Wiswall. 2016. "Testing the quantity–quality model of fertility: Estimation using unrestricted family size models." *Quantitative Economics*, 7(1): 157–192.

Myhrman, Antero, Paivi Olsen, Paula Rantakallio, and Esa Laara. 1995. “Does the wantedness of a pregnancy predict a child’s educational attainment?” *Family Planning Perspectives*, 27(3): 116–119.

Neal, Derek A., and William R. Johnson. 1996. “The role of premarket factors in black-white wage differences.” *Journal of Political Economy*, 104(5): 869–895.

Nguyen, Cuong Viet. 2018. “The long-term effects of mistimed pregnancy on childrens education and employment.” *Journal of Population Economics*, 31(3): 937–968.

North, A. Fredrick. 1979. “Health Services in Head Start.” In *Project Head Start: A Legacy of the War on Poverty.* , ed. Edward Zigler and Jeanette Valentine. New York:Free Press.

OEO. 1967. “3rd Annual Report: The Tide of Progress.”

OEO. 1970. “Annual Report Fiscal Years 1969-1970.”

Office of Child Development (DHEW). 1970. *Project Head Start 1968: A Descriptive Report of Programs and Participants.*

Ozbeklik, Serkan. 2014. “The Effect Of Abortion Legalization On Childbearing By Unwed Teenagers In Future Cohorts.” *Economic Inquiry*, 52(1): 100–115.

Pavan, Ronni. 2016. “On the production of skills and the birth order effect.” *Journal of Human Resources*, 51(3): 699–726.

Phillips, Deborah A., Mark W. Lipsey, Kenneth A. Dodge, Ron Haskins, Daphna Bassok, Margaret R. Burchinal, Greg J. Duncan, Mark Dynarski, Katherine A. Magnuson, and Christina Weiland. 2017. “The Current State of Scientific Knowledge on Pre-Kindergarten Effects.” *Accessed April 18, 2017.*

Pop-Eleches, Cristian. 2006. “The impact of an abortion ban on socioeconomic outcomes of children: Evidence from Romania.” *Journal of Political Economy*, 114(4): 744–773.

Potts, M, P Diggory, and J Peel. 1977. “Abortion Cambridge.” *UK: Cambridge University Press [Google Scholar].*

Price, Joseph. 2008. “Parent-child quality time: Does birth order matter?” *Journal of Human Resources*, 43(1): 240–265.

Rosenzweig, Mark R, and Kenneth I Wolpin. 1980. “Testing the quantity-quality fertility model: The use of twins as a natural experiment.” *Econometrica: journal of the Econometric Society*, 227–240.

Rosenzweig, Mark R, and Kenneth I Wolpin. 1993. “Maternal expectations and ex post rationalizations: the usefulness of survey information on the wantedness of children.” *Journal of Human Resources*, 23(2): 205–229.

Rosenzweig, Mark R, and Kenneth I Wolpin. 2000. "Natural" natural experiments" in economics." *Journal of Economic Literature*, 38(4): 827–874.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. "Integrated Public Use Microdata Series: Version 7.0 [dataset]." University of Minnesota.

Santelli, John, Roger Rochat, Kendra Hatfield-Timajchy, Brenda Colley Gilbert, Kathryn Curtis, Rebecca Cabral, Jennifer S Hirsch, and Laura Schieve. 2003. "The measurement and meaning of unintended pregnancy." *Perspectives on sexual and reproductive health*, 35(2): 94–101.

Schoen, Robert, Nan Marie Astone, Young J Kim, Constance A Nathanson, and Jason M Fields. 1999. "Do fertility intentions affect fertility behavior?" *Journal of Marriage and Family*, 61(3): 790–799.

Schultz, T Paul. 2005. "Effects of fertility decline on family well-being: Evaluation of population programs."

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. "What Are We Weighting For?" *Journal of Human Resources*, 50(2): 301–316.

Stuart, Bryan, Evan Taylor, and Martha Bailey. 2016. "Summary of procedure to match Numident place of birth county to GNIS places." U.S. Census Bureau 1284 Technical Memo 2.

Sulloway, Frank J. 2010. "Why siblings are like Darwins finches: Birth order, sibling competition, and adaptive divergence within the family." In *The Evolution of Personality and Individual Differences*. Oxford University Press New York.

Thompson, Owen. 2017. "Head Starts Long-Run Impact: Evidence from the Programs Introduction." *Journal of Human Resources*, 0216–7735r1.

United States Senate Committee on Labor and Public Welfare. 1964. "Economic Opportunity Act of 1964." U.S. Government Printing Office, Washington, DC.

Vinovskis, Maris. 2005. *The Birth of Head Start.* The University of Chicago Press.

Westinghouse Learning Corporation. 1969. *The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development. (Executive Summary).* Clearinghouse for Federal Scientific & Technical Information, Springfield, Va.

Westoff, Charles F, and Norman B Ryder. 1977. "The predictive validity of reproductive intentions." *Demography*, 14(4): 431–453.

Willis, Robert J. 1973. "A new approach to the economic theory of fertility behavior." *Journal of Political Economy*, 81(2, Part 2): S14–S64.