

The Child Left Behind: Parental Incarceration and Adult Human Capital in the US

Laura E. Henkhaus*

University of Southern California

December 31, 2018

[\[most recent version\]](#)

Abstract

Parental incarceration can be not only a huge economic shock to the family but also a source of psychological stress on family members. The effects of this stress may be long-lasting. Neuroscientists and psychiatrists describe how chronic stress in childhood may impair cognitive development. Exposure to parental incarceration is particularly prevalent in the United States, where over 8 percent of children have lived with a parent who was incarcerated during their childhood. In this paper, I investigate whether incarceration has long-term human capital consequences on children in the US. I provide evidence at the population level that parental incarceration causes lower rates of high school diploma receipt and likely causes lower rates of full-time employment in young adulthood. This work adds to the body of evidence documenting an intergenerational transmission of socioeconomic disadvantage and has important implications for social policy. Within the education system, results might motivate improved support for children's socio-emotional health.

JEL Classification: I24, I30, J15, J2

* I thank Robynn Cox, Darius Lakdawalla, and John Romley for helpful comments. I conducted this work with generous support from a Schaeffer-Amgen Predoctoral Fellowship and the USC School of Pharmacy. This research uses data from Add Health, a program directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from Grant P01-HD31921 for this analysis.

Email: laura.henkhaus@usc.edu

1 Introduction

The United States is known to have the highest incarceration rate in the world (Wagner and Sawyer 2018).¹ An accompanying fact is that many US children have grown up with a parent incarcerated. By ages 6 to 17, 8 percent of US children have lived with a parent who had been incarcerated during their lifetime (Murphey and Cooper 2015). The impact of parental incarceration on children certainly varies by case, and effects at the population level are ambiguous. Removal of an abusive parent from the home would improve the child's health in at least one respect but still may be a distressing process. Parental incarceration, however, almost always comes with huge economic cost to families. Beyond loss in income, families also face large financial costs from court fines, inmate room and board, and visitation fees (Cox 2018, Grinstead et al. 2018). Extant research documents hardship faced by children with a parent incarcerated, who face increased risks of household food insecurity (Cox and Wallace 2016), homelessness (Wildeman 2014), and school expulsion or suspension (Johnson 2009).

The psychological and economic effects of parental incarceration may have lasting impacts on educational attainment and labor market outcomes in adulthood. Distress from missing the incarcerated parent, from stigma of having an incarcerated parent, or from unmet physical needs may be barriers to focus and learning in school. Parental incarceration may increase children's likelihood to participate in crime due to economic need or through a role model effect, thereby limiting employment opportunities. Economic strain on families may squeeze investment in children, as families face trade-offs between costs to maintain

¹ Among nations with populations of at least 500,000.

contact with the incarcerated family member and costs to support children through counseling and academic enrichment, for example.

The main challenge to measuring effects of parental incarceration is that the groups exposed and not exposed are very different. For example, parents in households where one has been incarcerated have lower educational and occupational attainment (Mears and Siennick 2016). And, incarceration rates are much higher among Blacks than among Whites or Hispanics (Cox 2018).

In the present study, I use nationally representative US data to evaluate whether parental incarceration impacts children's educational attainment and labor market outcomes in young adulthood. I examine robustness of results to varying assumptions about remaining selection on unobservables, using information on selection on observables. My work is distinct from studies of the effects of incarceration on family members which use administrative data in jurisdictions requiring random assignment of cases to judges.² Such work measures the effects of incarceration of only the marginal accused person, for whom the judge's propensity to deliver a sentence of jail or prison time is believed to be the deciding factor of whether the accused is incarcerated or not.

² For example, see Aizer and Doyle Jr. (2015) in the context of Chicago, IL on the effects of own incarceration during childhood on human capital; Dobbie, Grönqvist, Niknami, Palme and Priks (2018) in the context of Sweden on the effects of parental incarceration on teen pregnancy, teen crime, and employment in early adulthood; and Bhuller, Dahl, Loken and Mogstad (2018) in the context of Norway on the effects of parental incarceration on school grades and being charged with a crime.

2 Data

I utilize restricted-use data files from the National Longitudinal Study of Adolescent to Adult Health (Add Health), which recruited a random sample of children in grades 7 to 12 (Harris 2013). Recruitment occurred first by sampling secondary schools from a national database, then by sampling students within the schools. The sample includes 132 schools—a mix of public and private schools (Harris et al. 2006). Some groups were oversampled, such as minority racial groups, disabled adolescents, and siblings (Harris 2013). Thus, I use survey weights in analyses so that results reflect a nationally representative sample.

I use data on participants from Waves I (collected 1994-1995 when participants were aged 11-18 years old) and Wave IV (2008-2009, 24-32 years old³) as well as data from Wave II (1996, 12-19 years old) and Wave III (2001-2002, 18-26 years old) to replace missing data where possible and appropriate. Parents were invited to complete a separate questionnaire during Wave I. The full sample which completed both Wave I and Wave IV interviews consists of roughly 15,000 individuals. The unweighted response rate for Wave IV was 80 percent. Analyses from Add Health study staff indicated that nonresponse bias was negligible and that participants in Wave IV were representative of the original cohort recruited in Wave I (Brownstein et al. 2011).

³ Fifty-two respondents were aged 33-34 years old at time of Wave IV interview.

2.1 Key measures

2.1.1 Parental incarceration

I define parental incarceration as report in Wave IV of either biological parent, mother figure, or father figure being incarcerated while the child was alive but not yet 18-years-old. In constructing the parental incarceration measure, I categorized responses of “don’t know” as such rather than designating these values as missing. The “don’t know” group consists of (i) respondents selecting “don’t know” in the initial question of whether the parent had been incarcerated, which was largely due to the parent’s absence in the respondent’s life⁴ and (ii) respondents who acknowledged that a parent had been incarcerated but responded “don’t know” to questions on age when parent was first incarcerated and last released.

2.1.2 Childhood abuse

I consider three types of childhood abuse: sexual, physical, and emotional abuse. I define the sample of adults who experienced childhood sexual abuse as those who reported either that they experienced nonconsensual sexual touching or sexual relations by an adult caregiver before sixth grade (question in Wave III) or before age 18 (Wave IV) or forced sexual activity by a non-caregiver before age 18 (Wave IV). Childhood physical abuse reflects report of at least one time to the question, “Before your 18th birthday, how often did a parent or adult caregiver hit you with a fist, kick you, or throw you down on the floor, into a wall, or down stairs?” Emotional abuse reflects the maximum count, at least ten times, to the question:

⁴ The majority of this group also responded “don’t know” to the preceding question of whether the parent was still alive.

“Before your 18th birthday, how often did a parent or other adult caregiver say things that really hurt your feelings or made you feel like you were not wanted or loved?”

A key feature of Add Health is that, while the survey was largely administered through computer-assisted personal interview, questions on sensitive topics such as childhood mistreatment were completed through audio computer-assisted *self-interview*. Self-interview methods have been found to capture higher rates of sexual and drug-related behaviors than measured from face-to-face interviews (Midanik and Greenfield 2008, Perlis et al. 2004). All questions used to construct the childhood abuse measures were part of the self-interview section.

2.1.3 Neighborhood and family background

I use identifiers for the secondary schools that participants attended and measures of socioeconomic status of the family level. The covariates representing family SES include parent-reported household income and highest parental educational attainment.

2.1.4 Human capital outcomes

I study young adult outcomes reported in Wave IV: having a high school diploma, a college degree, employment, full-time employment, and earnings level. A person is employed if currently working at least 10 hours per week. Full-time employment status is defined as working at least 35 hours per week. Earnings include any wages, bonuses, and self-employment.

2.2 Missing data and imputation

Because this research relies on longitudinal survey data, it is important to examine the extent and nature of missing data. The rates of missing data were generally low except for one key variable only available for children whose parents participated in interviews: household income. I used multiple imputation for analysis, as described below.

The sample who completed Wave I and Wave IV interviews, with survey weights calculated by Add Health, consists of 15,642 people. The rates of missing data for all but two variables in the fully controlled models used here are less than 2 percent. The rate missing for adult earnings is almost 5 percent, and the rate missing for childhood household income is 24 percent. Wave I parent interviews were completed for not all but over 85 percent of adolescents who participated in the Wave I in-home interviews (Harris 2013).

I created an analytic file with missing observations imputed and conducted analysis on the multiply imputed data. The multiple imputation procedure consists of three steps: (i) impute missing values m times to create m data sets, (ii) execute analysis models on each of the m data sets, (iii) use Rubin's rules to combine the estimates produced from each of the m data sets in the prior step (Rubin 1996, Rubin 1987). I imputed 25 sets of values for childhood household income.⁵ The imputation procedure uses the distribution of observed data to impute multiple values. Each imputed value includes a random component, reflecting the uncertainty around the true value (Rubin 1987, Schomaker and Heumann 2018). The variance of a coefficient estimate obtained through multiple imputation is calculated as the

⁵ White, Royston and Wood (2011) suggest a rule of thumb that the number of imputations should be at least as large as the largest fraction of missing data * 100, across variables. Here, the largest fraction of missing data is 0.24, for log-household income. Results from baseline models on data imputed 50 times were identical or nearly identical to results pooled from the 25 imputations.

sum of variance within (mean variance from the individual imputations), variance between (variance across coefficient estimates from the m imputations), and additional sampling variance considering the number of imputations m (variance between / m) (Rubin 1996). See the Appendix for more details on the imputation procedure.

2.3 Sample characteristics

In this nationally representative sample, 11.7 (0.6) percent of adults reported having experienced at least part of childhood with a parent incarcerated. Those exposed to parental incarceration experienced other types of childhood adversity at higher rates. They were twice as likely to experience each (not all) type of childhood abuse. Median household income for children who experienced parental incarceration was about 55 percent of median income of households in which no child experienced parental incarceration. There were stark differences by race. Among the sample exposed to parental incarceration in childhood, the percent of respondents who were Black non-Hispanic was nearly double the rate in the group never exposed to parental incarceration in childhood. Children who experienced parental incarceration had much higher rates of depressive symptoms but no difference in rate of educational challenges (learning disability or ADHD). The large disparity in rates of depressive symptoms persisted in adulthood, when those who had been exposed to parental incarceration had lower educational attainment and worse labor market outcomes. See descriptive statistics in Table 1 below.

Table 1. Sample means across parental incarceration history^a
(standard errors in parentheses)

	No parental incarceration	Parental incarceration
	N=12,317	N=1,893
Female	0.491 (0.008)	0.504 (0.018)
Race/ethnicity		
White NH	0.789 (0.023)	0.670 (0.036)
Black NH	0.133 (0.018)	0.251 (0.036)
Native American NH	0.008 (0.002)	0.019 (0.01)
Asian/Pacific Islander NH	0.032 (0.008)	0.007 (0.003)
Other NH	0.032 (0.003)	0.049 (0.01)
Hispanic	0.109 (0.017)	0.137 (0.021)
Childhood household income, median (2010\$)	\$61,721	\$33,800
Child health and cognition		
Child depressive symptoms	0.289 (0.008)	0.387 (0.018)
Educational challenge	0.157 (0.012)	0.168 (0.014)
Adverse childhood experiences		
Parental incarceration	0.000	1.000
Sexual abuse	0.120 (0.005)	0.240 (0.014)
Physical abuse	0.149 (0.006)	0.340 (0.018)
Emotional abuse	0.101 (0.004)	0.226 (0.014)
Adult outcomes		
Adult depressive symptoms	0.261 (0.008)	0.376 (0.016)
High school diploma	0.873 (0.009)	0.682 (0.017)
College degree	0.352 (0.017)	0.124 (0.012)
Full-time employment	0.737 (0.008)	0.652 (0.016)
Adult earnings, median (2010\$)	\$30,414	\$22,304

Key. NH, Non-Hispanic.

^a Except where noted as "median."

* p<0.05, ** p<0.01, *** p<0.001

3 Empirical strategy

I employ multiple strategies to evaluate whether there is a causal impact of parental incarceration on education or labor market outcomes. First, in the baseline specification I implement school fixed effects regression models controlling for family-level SES and other adverse childhood experiences (ACEs), among other confounders. Second, I examine robustness of results by constructing bounds under varying assumptions about the importance of unobservable confounders. Third, I directly test for selection with two sets of falsification tests.

3.1 Baseline model

In my baseline approach to study the effects of parental incarceration on adult human capital outcomes, I use a school fixed effects strategy while controlling for demographics, childhood socioeconomic status, and other adverse childhood experiences. In the baseline specification below, let Y_i denote the outcome for individual i . Y_i is either a binary variable for having a high school diploma, having a college degree, currently employed (at least 10 hours/week), currently employed full-time, or a continuous measure of earnings level. The parameter of interest is α in the following equation:

$$Y_i = \alpha \cdot \text{ParentalIncarceration}_i + X_i' \beta + \varepsilon_i \quad (1)$$

$\text{ParentalIncarceration}_i$ is a dummy variable equal to one for persons who reported having a parent incarcerated at any time during ages 0-17. X_i is a vector of controls and includes a constant. Other components of X_i include demographics (age when outcome measured, sex, and race/ethnicity), childhood household socioeconomic status (highest

parental education level and household income), school fixed effects, and other adverse childhood experiences (sexual abuse, physical abuse, emotional abuse, and “don’t know” if parent incarcerated). Lastly, the model allows an individual, idiosyncratic error term ε_i .

By including school fixed effects, I am able to parse out unobserved characteristics of the school and school neighborhood environments that affect both the likelihood of parental incarceration and the likelihood of success in school and on the labor market. For example, descriptive statistics show that a child in a low-income neighborhood (not shown) is more likely to experience parental incarceration. And, we know that schools in low-income neighborhoods, on average, provide lower quality education than is available to children who attend schools in higher income neighborhoods. Interpreting these baseline results causally requires the assumption that there is no unobserved heterogeneity across children at the same schools which influence both their likelihood of parental incarceration and their education and labor market outcomes. In the following sections, I address the possibility of remaining unobserved confounding.

I use ordinary least squares (OLS) to model binary outcomes and replicate these models with logit and probit regression in ancillary analyses. I analyze annual earnings with a two-part model (Dow and Norton 2003, Roodman 2009) to include people with zero earnings. The first part is a probit model estimating the probability of having positive earnings, and the second part is a generalized linear model (log link function and gamma distribution) estimating earnings levels, using observations with positive earnings.

3.2 Robustness checks

3.2.1 Effect bounds: quantifying selection on unobservables

In this section, I calculate bounds on the effects of parental incarceration under varying assumptions on the degree of selection on unobservables relative to selection on observables, following partial identification methods developed by Altonji, Elder, and Taber (2002, 2005) and Oster (2017). Partial identification allows the researcher to recover bounds on the estimated treatment effects in contexts where unobservable variables may cause confounding. This method relies on the premise that confounding on observables provides insight into the influence of confounding on unobservables.

In this context, I assessed whether unobserved factors, such as low parental investment in children, fully explain the results. I adopted Oster's approach for linear models (Oster 2017). This method involves using observable confounders to bound the likely effect of unobservable confounders. Thus, key parameters are: (i) the amount of selection on unobservables relative to selection on observables and (ii) the amount of outcome variance that would be explained by the unobservable confounders. The first value, the selection parameter, is denoted δ below, and the second value informs the value R_{max} , which represents the R^2 from the hypothetical regression including unobservable confounders on the right-hand-side. I construct effect bounds for varying values of δ from 0 to 1. I make the following alternative assumptions about the value R_{max} :

(i) $R_{max} = 1.3\tilde{R}$, where \tilde{R} is the R^2 from equation (1) above with the full set of observed controls, and

(ii) $R_{max} = 2\tilde{R}$.

Below, I describe the approach to construct bounds on the effects of parental incarceration. To start, consider a modified version of equation (1) above, omitting the individual subscripts i :

$$Y = \tilde{\alpha} \cdot \text{ParentalIncarceration} + X' \tilde{\beta} + W_2 + \eta \quad (3)$$

The new term, W_2 , represents unobservables which determine the human capital outcome Y and are correlated with parental incarceration but not correlated with any of the observable confounders. For example, unobserved level of parental investment in children will be contained in W_2 if the partial correlation with parental incarceration, conditional on the observed controls, is non-zero. By definition as a confounder, $\text{cov}(W_2, Y) \neq 0$ and $\text{cov}(W_2, \text{ParentalIncarceration}) \neq 0$. An additional requirement of W_2 is that $\text{cov}(W_2, W_1) = 0$, denoting $W_1 = X' \tilde{\beta}$. This orthogonality requirement implies that W_1 captures observables in addition to any confounding from unobservables which are correlated with the observables. Now, the parameter on parental incarceration, $\tilde{\alpha}$, is the true effect on outcome Y . The goal is to estimate bounds for $\tilde{\alpha}$, but the coefficients on the observed variables in W_1 should not be interpreted causally. Recall that X includes sexual, physical, and emotional abuse along with an indicator for the group designated as “don’t know” whether a parent was incarcerated during childhood (part of this group acknowledged parental incarceration but responded “don’t know” to questions on own age when parent was incarcerated), Thus while elements of W_2 are correlated with parental incarceration, they must be uncorrelated with these other types of childhood adversity. Lastly, η is the individual idiosyncratic error term which contains unobservables which determine the outcome Y but are uncorrelated with parental incarceration, controls, or W_2 .

The ratio of selection on unobservables to selection on observables is defined as

$$\delta = \frac{\sigma_{2,p}}{\sigma_2^2} / \frac{\sigma_{1,p}}{\sigma_1^2}$$

where $\sigma_{j,p} = \text{cov}(W_j, \text{ParentalIncarceration})$ and $\sigma_j^2 = \text{var}(W_j)$ for $j \in \{1, 2\}$. The denominator of this ratio—the level of selection on observables—can be readily computed from the data. The numerator—the selection on unobservables—is identified with a restriction on the parameter $\tilde{\alpha}$, the effect of parental incarceration, and a restriction on R_{max} . Here, I choose to compute bounds on the effect $\tilde{\alpha}$, thus I assume various values for the selection ratio δ and R_{max} .

I calculate bounds on effect sizes for a set of R_{max} values informed by the R^2 from the regression with the full set of observed controls. In most cases, it is unlikely that $R_{max} = 1$, *i.e.*, that the treatment/exposure, controls, and the unobservable component W_2 fully explain the outcome, due to measurement error in the outcome or other idiosyncratic variation in the outcome (Oster 2017). In the present case, there are likely important determinants of human capital outcomes which were determined after Wave I but not caused by these childhood factors. For example, all cases of depression not caused by parental incarceration (or controls such as the other adverse childhood experiences) will be included in an individual's η error term.

The first R_{max} condition above, $R_{max} = 1.3\tilde{R}$, was proposed by Oster from calculations with data from randomized studies. From her set of results from randomized studies published in top economics journals which reported uncontrolled and controlled estimates, she calculated that when holding fixed $\delta=1$, 90 percent of the results would survive if $R_{max} =$

$1.3\tilde{R}$ (Oster 2017).⁶ I present the main results of the bounding exercise under this condition: $R_{max} = 1.3\tilde{R}$. Note that the second R_{max} condition, $R_{max} = 2\tilde{R}$, implies that the unobserved confounders explain as much variance in the outcome as do the observables.

Interpreting the ratio δ here is difficult given the large amount of selection on observables. In general, if the estimated treatment (or exposure) effect erodes to zero only when $\delta > 1$, this is considered evidence that at least part of the estimated effect is real (Altonji, Elder, and Taber 2005, Altonji, Elder, and Taber 2002, Oster 2017). However, the true ratio of selection on unobservables to selection on observables, δ , is likely less than 1 and may be much less than 1, in particular when selection on observables is substantial as in the present case. I calculate bounds for the effect size under the conditions $\delta=0, 0.25, 0.5, 1$.

I calculate bounds on the effects of parental incarceration on the following outcomes: high school diploma receipt, college degree attainment, and full-time employment. I do not compute bounds for log earnings because results from equation (1) showed no effect of parental incarceration on earnings in the sample with positive earnings.

3.2.2 Falsification tests

I implement direct tests of selection in two ways. In the first set of falsification tests, I evaluate whether worse outcomes were also observed among those who had a parent who was incarcerated but not during their childhood, which would suggest that results might be driven by unobserved family qualities rather than direct experience of exposure to parental

⁶ To “survive” means both that the identified set does not include zero and that the identified set is within 2.8 standard errors of the fully controlled estimate. The sample of results selected in Oster (2017) come from randomized studies in top-five economics journals over a six-year period. Across this set of 65 results, she found that 90 percent would survive a cutoff of $R_{max} = 1.3\tilde{R}$, and only 40 percent would survive a cutoff of $R_{max} = 1$.

incarceration. As a second test, I examine whether there was a pre-existing difference in cognitive ability by exploiting child age when parent was incarcerated and the timing of the Wave I survey when a picture vocabulary test was administered.

I use data on the participant's age when parent was first incarcerated and when last released. I construct five groups: (i) parent incarcerated and last released before child was born, (ii) parent incarcerated after child's birth but before age 18, (iii) parent first incarcerated after child turned 18, (iv) parent never incarcerated—the reference group, and (v) “don't know,” reflecting either a “don't know” response for timing of parent's incarceration or “don't know” response to whether a parent had ever been incarcerated. If either group (i) or (iii) have poorer outcomes compared to the group whose parents were never incarcerated, conditional on the full control set, then there is evidence that it may be unobserved family qualities and not direct exposure to parental incarceration that cause poorer adult human capital outcomes.

4 Results

4.1 Baseline results

Baseline results show that adults who had been exposed to parental incarceration in childhood had worse education and labor market outcomes. The unadjusted differences in means for the education variables were large: -18.5 percentage points for high school diploma and -23.6 percentage points for college degree. The difference in unadjusted mean for full-time work was -8.1 percentage points and for earnings was -\$9,766, or 32.1 percent of median earnings for this young adult sample. In the fully adjusted models, the differences

in educational attainment and labor market outcomes persist but do reduce substantially. The regressions with the full control set show that the estimates of the effect of parental incarceration were -10.0 percentage points for high school diploma attainment, -7.8 percentage points for college degree attainment, -4.3 percentage points for full-time work status, and -\$3,645 for earnings (12.0 percent of median earnings). For each outcome, more than 55 percent of the reduction in average marginal effects for parental incarceration were explained by household SES. However, this was not just due to parental incarceration causing lower household income. Sensitivity analyses omitting household income from control set (3)—that is, controlling for household SES only through highest parental education—show that parental education contributes to roughly half of the reduction between the uncontrolled and fully controlled estimates for parental incarceration. See results in Table 2.

Table 2. Estimates of the effects of parental incarceration on education and labor market outcomes

Average marginal effects (standard errors in parentheses)					
Controls	None ^a	Demographics ^b	Col. 2 + household SES ^c	Col. 3 + school FE ^d	Col. 4 + other ACEs ^e
	(1)	(2)	(3)	(4)	(5)
A. High school diploma; mean (s.d.): 0.828 (0.378)					
OLS	-0.185*** (0.017)	-0.176*** (0.018)	-0.122*** (0.019)	-0.109*** (0.018)	-0.100*** (0.017)
<i>R</i> ²	0.034	0.044	0.101	0.153	0.155
Adjusted <i>R</i> ²	0.034	0.044	0.100	0.144	0.146
B. College degree; mean (s.d.): 0.302 (0.459)					
OLS	-0.236*** (0.017)	-0.217*** (0.017)	-0.106*** (0.013)	-0.092*** (0.014)	-0.078*** (0.013)
<i>R</i> ²	0.032	0.055	0.209	0.259	0.264
Adjusted <i>R</i> ²	0.032	0.055	0.208	0.251	0.256
C. Employment; mean (s.d.): 0.805 (0.396)					
OLS	-0.063***	-0.059***	-0.037*	-0.039*	-0.030*

	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)
R^2	0.005	0.023	0.033	0.076	0.068
Adjusted R^2	0.005	0.023	0.032	0.067	0.058
D. Full-time employment; mean (s.d.): 0.706 (0.455)					
OLS	-0.081***	-0.076***	-0.055***	-0.052**	-0.043*
	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)
R^2	0.006	0.039	0.045	0.097	0.093
Adjusted R^2	0.005	0.038	0.044	0.088	0.083
E. Earnings; mean (s.d.): \$35,025 (44,170); median: \$30,414					
2PM	-\$9,766***	-\$8,450***	-\$4,262***	-\$3,766***	-\$3,632**
	(1,318)	(1,230)	(1,221)	(1,109)	(1,101)

NOTE. — Sample size is N=14,741 for each regression here.

^a The only right-hand side variables are parent incarcerated during childhood and “don’t know” if parent incarcerated when aged 0-17 years.

^b Demographic controls include age when outcome measured; sex as female or male; and race/ethnicity as White Non-Hispanic, Black Non-Hispanic, Native American Non-Hispanic, Asian/Pacific Islander Non-Hispanic, Other Non-Hispanic (includes multi-racial), or Hispanic.

^c Household SES controls include log of childhood household income and highest parental educational attainment as (i) less than high school, (ii) GED, (iii) high school diploma, (iv) vocational school after high school, (v) some college, (vi) college graduate, or (vii) beyond 4-yr college.

^d School fixed effects: schools were the primary sampling unit.

^e Other ACEs include childhood sexual, physical, and emotional abuse.

* p<0.05, ** p<0.01, *** p<0.001

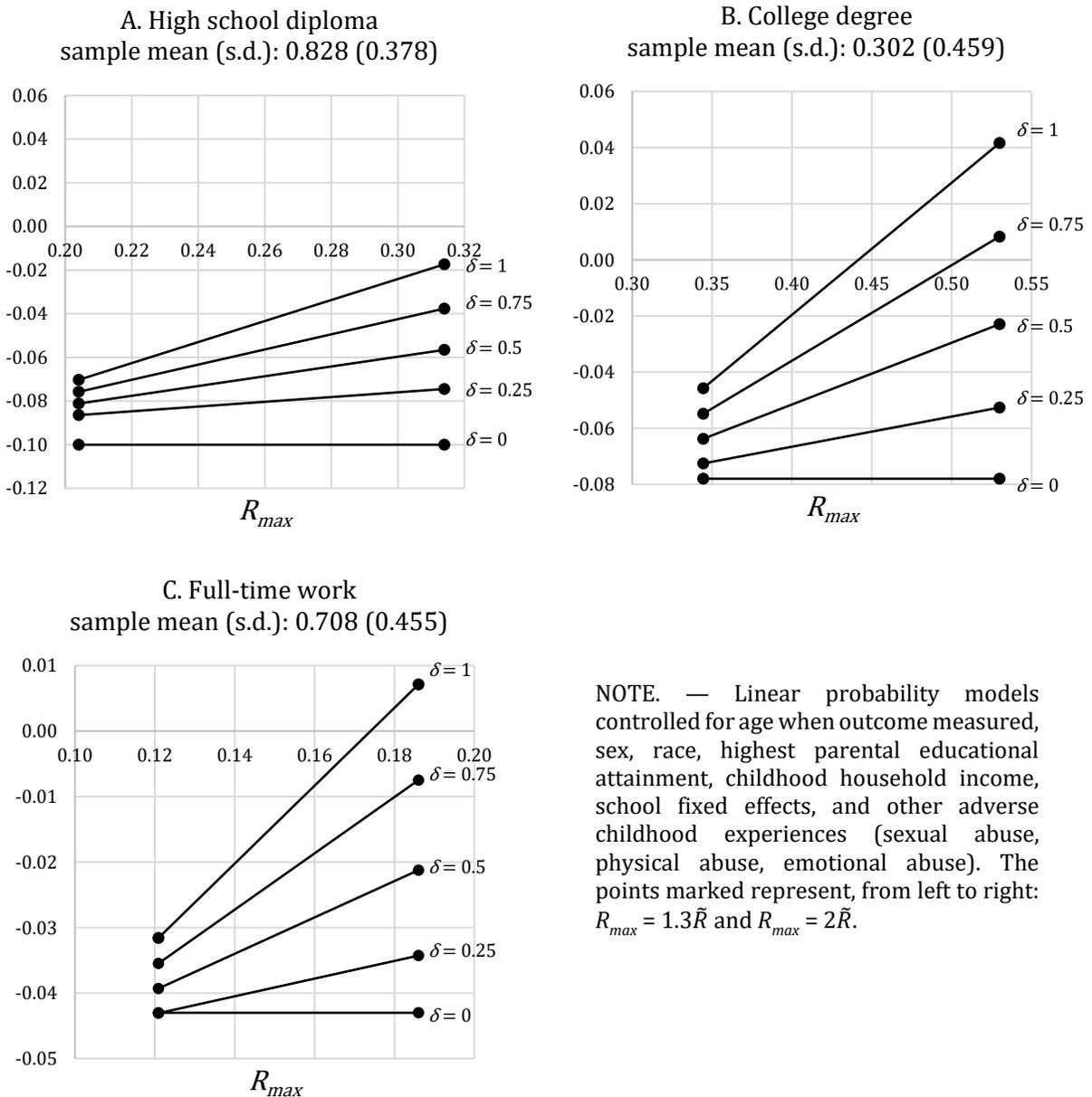
4.2 Robustness checks

4.2.1 Effect bounds: quantifying selection on unobservables

Evaluation of robustness of results to varying assumptions about remaining selection on unobservables provide support for real, negative effects of parental incarceration on high school diploma receipt and full-time employment. Under the condition $R_{max} = 1.3\tilde{R}$, the estimates remain negative for all values of δ . The lower bound for the effect of parental incarceration on high school diploma receipt (when $\delta=1$) is -7.1 percent. A negative estimate persists at a higher value of $R_{max} = 2\tilde{R}$ for all values δ until $\delta = 1$. Studying full-time work,

estimates remain negative under the condition $R_{max} = 2\tilde{R}$ and $\delta = 1$. Results for college degree attainment were less robust to the condition $R_{max} = 2\tilde{R}$. See Figure 1 below.

Figure 1. Bounds on the effects of parental incarceration on outcomes of adult children



4.2.2 Falsification tests

Falsification tests were successful—showing that parental incarceration occurring only before the child was born or only after the child became an adult did not affect the outcomes.

See results in Table 3 below.

Table 3. Falsification test: parental incarceration

Average marginal effects (s.e.)

	Outcome			
	High school diploma	College degree	Full-time employment	Earnings
Timing of parental incarceration				
Before birth	-0.034 (0.044)	-0.020 (0.047)	-0.016 (0.058)	\$1,128 (2,843)
Age 0 - 17	-0.102 (0.017)***	-0.075 (0.014)***	-0.039 (0.017)*	-\$3,810 (1,147)
Age 18+	-0.058 (0.036)	-0.031 (0.031)	0.041 (0.031)	-\$2,999 (2,221)
“Don’t know” ^a	-0.073 (0.018)***	-0.048 (0.016)**	-0.037 (0.021)	-\$3,329 (1,579)
Never	ref.	ref.	ref.	ref.
N	13,563	13,503	13,560	13,041

NOTE. — The table shows the estimated average marginal effects and standard errors for the regressions corresponding to equation (1), including the full control set (OLS for binary outcomes, two-part model for earnings).

^a “Don’t know” indicates response of “don’t know” either to the questions on age when parent was first incarcerated and last released (among those acknowledging parental incarceration) or to any question on whether biological parents or parental figure (when not the biological parent) had ever been incarcerated.

* p<0.05, ** p<0.01, *** p<0.001

And, conditional on the control set, there was no pre-existing difference in childhood vocabulary scores.

5 Conclusion

This paper provides, to my knowledge, the first causal evidence at the US population level of negative effects of parental incarceration on children's educational attainment and likelihood of employment. This work adds to the body of evidence documenting an intergenerational transmission of socioeconomic disadvantage and has important implications for the criminal justice system. Within the education system, results might motivate improved support for children's socio-emotional health.

References

- Aizer, Anna, and Joseph J. Doyle Jr.** 2015. "Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges." *Quarterly Journal of Economics* 130 (2):759-803.
- Altonji, Joseph G, Todd E Elder, and Christopher R Taber.** 2002. Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. Dept. Econ., Northwestern University.
- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber.** 2005. "Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools." *Journal of Political Economy* 113 (1):151-184.
- Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Loken, and Magne Mogstad.** 2018. "Intergenerational effects of incarceration." *AEA Papers and Proceedings* 108:234-40.
- Brownstein, Naomi, William D Kalsbeek, Joyce Tabor, Pamela Entzel, Eric Daza, and Kathleen Mullan Harris.** 2011. "Non-response in wave IV of the National Longitudinal Study of Adolescent Health." *Carolina Population Center, University of North Carolina: Chapel Hill.*
- Cox, Robynn.** 2018. "Mass Incarceration, Racial Disparities in Health, and Successful Aging." *Generations* 42 (2):48-55.
- Cox, Robynn, and Sally Wallace.** 2016. "Identifying the link between food security and incarceration." *Southern Economic Journal* 82 (4):1062-1077.
- Dobbie, Will, Hans Grönqvist, Susan Niknami, Mårten Palme, and Mikael Priks.** 2018. "The Intergenerational Effects of Parental Incarceration." *National Bureau of Economic Research Working Paper Series* No. 24186.
- Dow, William H., and Edward C. Norton.** 2003. "Choosing between and interpreting the Heckit and two-part models for corner solutions." *Health Services and Outcomes Research Methodology* 4 (1):5-18.
- Grinstead, Olga, Bonnie Faigeles, Carrie Bancroft, and Barry Zack.** 2018. "The Financial Cost of Maintaining Relationships with Incarcerated African American Men: Results from a Survey of Women Prison Visitors." In *Impacts of Incarceration on the African American Family*, 61-70. Routledge.
- Harel, Ofer.** 2009. "The estimation of R^2 and adjusted R^2 in incomplete data sets using multiple imputation." *Journal of Applied Statistics* 36 (10):1109-1118.
- Harris, K. M., C. T. Halpern, A. Smolen, and B. C. Haberstick.** 2006. "The National Longitudinal Study of Adolescent Health (Add Health) twin data." *Twin Res Hum Genet* 9 (6):988-97.
- Harris, Kathleen Mullan.** 2013. The Add Health Study: design and accomplishments. Carolina Population Center, University of North Carolina at Chapel Hill.
- Johnson, Rucker.** 2009. "Ever-increasing levels of parental incarceration and the consequences for children." In *Do prisons make us safer? The benefits and costs of the prison boom*, edited by Steven Raphael and Michael A. Stoll, 177-206.
- Mears, Daniel P., and Sonja E. Siennick.** 2016. "Young Adult Outcomes and the Life-Course Penalties of Parental Incarceration." *Journal of Research in Crime and Delinquency* 53 (1):3-35.

- Midanik, L. T., and T. K. Greenfield.** 2008. "Interactive voice response versus computer-assisted telephone interviewing (CATI) surveys and sensitive questions: the 2005 National Alcohol Survey." *J Stud Alcohol Drugs* 69 (4):580-8.
- Murphey, David, and P. Mae Cooper.** 2015. *Parents behind bars: What happens to their children.* Bethesda, MD: Child Trends.
- Nguyen, Cattram D., John B. Carlin, and Katherine J. Lee.** 2017. "Model checking in multiple imputation: an overview and case study." *Emerging Themes in Epidemiology* 14 (1):8.
- Oster, Emily.** 2017. "Unobservable selection and coefficient stability: theory and evidence." *Journal of Business & Economic Statistics*:1-18.
- Perlis, Theresa E, Don C Des Jarlais, Samuel R Friedman, Kamyar Arasteh, and Charles F Turner.** 2004. "Audio-computerized self-interviewing versus face-to-face interviewing for research data collection at drug abuse treatment programs." *Addiction* 99 (7):885-896.
- Roodman, David.** 2009. "Estimating fully observed recursive mixed-process models with cmp."
- Rubin, Donald B.** 1987. *Multiple imputation for nonresponse in surveys.* New York: Wiley.
- Rubin, Donald B.** 1996. "Multiple imputation after 18+ years." *Journal of the American Statistical Association* 91 (434):473-489.
- Schomaker, Michael, and Christian Heumann.** 2018. "Bootstrap inference when using multiple imputation." *Statistics in Medicine* 37 (14):2252-2266.
- Wagner, Peter, and Wendy Sawyer.** 2018. *States of Incarceration: The Global Context 2018.* Northampton, MA: Prison Policy Initiative.
- White, I. R., P. Royston, and A. M. Wood.** 2011. "Multiple imputation using chained equations: issues and guidance for practice." *Statistics in Medicine* 30 (4):377-99.
- Wildeman, Christopher.** 2014. "Parental incarceration, child homelessness, and the invisible consequences of mass imprisonment." *The Annals of the American Academy of Political and Social Science* 651 (1):74-96.

Appendix

Multiple imputation

In all regression analyses, I addressed missing data by implementing multiple imputation. I report pooled regression results from the multiply imputed data sets, adjusting standard errors to account for the imputation procedure.

Missing values for childhood household income were imputed using chained equations. Each of these equations included survey weights. I imputed values for childhood family income as well as for vocabulary score for the falsification check. In the imputation model, I included the covariates from the analysis models (control set described above) as well as auxiliary variables correlated with the missing variables. I included variables from the analytic model to preserve the relationships between the variables of interest (White, Royston, and Wood 2011, Nguyen, Carlin, and Lee 2017, Rubin 1987). As auxiliary variables in the imputation model, I included variables for full-time work status of the resident father and median household income of the Census block group in which the child lived during Wave I. I considered the following sociodemographic factors as auxiliary variables as well but excluded from the imputation model due to low correlation with the variables for which values were to be imputed (correlation < 0.3): full-time work status of resident mother, parental respondent age, parental respondent race, and parental respondent US-nativity.

Variance of parameters estimated through multiple imputation is the sum of three components: variance within, variance between, and an additional source of sampling variance. Variance within is the arithmetic mean across sampling variances for an estimate across each of the imputed data sets—i.e., the sampling variability expected had there been no missing data. Variance between is the variance of parameter estimates across each imputed data set, *i.e.*, it captures additional uncertainty that arises from missing data. Lastly, the additional sampling variance is the variance between divided by the number of imputations, representing sampling error of the average coefficient estimates, which is larger for smaller number of imputations. Thus, the standard errors for each parameter estimate obtained through multiple imputation is the square root of this total variance across the three components (Rubin 1987).

Average marginal effects from the nonlinear models and R^2 values from the OLS models were also calculated using Rubin's rules (Rubin 1996). Due to the multiple imputation strategy, the R^2 values are not from one model but rather represent the average R^2 across all imputations, using Fisher's transformation. Following Harel (2009), the R^2 from each imputation was transformed to a correlation, then Fisher's transformation was used to convert the correlation to a z-score. The mean z-score was calculated across imputations then transformed back to an R^2 value. A simulation study by Harel suggests that while R^2 values obtained through this transformation procedure tend to be biased upward, adjusted R^2 values tend to be biased downward (Harel 2009). Thus, both R^2 measures are reported in the paper.