CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Early Education and Adult Health: Age 37 Impacts and Economic Benefits of the Child-Parent Center Preschool Program

Nishank Varshney[1,2,*] ![ORCID], Judy A. Temple[1,2] and Arthur J. Reynolds[2,3]

[1]Humphrey School of Public Affairs, University of Minnesota, 301 19th Avenue South, Minneapolis, MN 55455, United States
[2]Human Capital Research Collaborative, University of Minnesota, 51 E River Road, Minneapolis, MN 55455, United States
[3]Institute of Child Development, University of Minnesota, 51 E River Road, Minneapolis, MN 55455, United States
*Corresponding author: E-mail: varsh011@umn.edu

Abstract
This article evaluates the long-term impacts of the Chicago Child-Parent Centers (CPC), a comprehensive early childhood program launched in the 1960s, on physical and mental health outcomes. This study follows a cohort of 1539 participants born in 1979–1980 and surveyed most recently at age 35–37 by employing a matched study design that included all 989 children who entered CPCs at ages 3 and 4 (1983-1985) and 550 comparison children of the same age from randomly selected schools participating in the usual district early childhood programs in kindergarten. Using propensity score weighting that addresses potential issues with differential attrition and non-random treatment assignment, results reveal that CPC preschool participation is associated with significantly lower rates of adverse health outcomes such as smoking and diabetes. Further, evaluating the economic impacts of the preschool component of the program, the study finds a benefit-cost ratio in the range of 1.35–3.66 (net benefit: $3896) indicating that the health benefits of the program by themselves offset the costs of the program even without considering additional benefits arising from increased educational attainment and reduced involvement in crime reported in earlier cost-benefit analyses. The findings are robust to corrections for multiple hypothesis testing, sensitivity analysis using a range of discount rates, and Monte Carlo analysis to account for uncertainty in outcomes.

1 Introduction

Evidence continues to accumulate on the significance of early childhood education in improving children's long-term development and well-being. Studies in neuroscience, education, and social science indicate that improvements in early education can have

meaningful impacts on the lives of children (Heckman, 2008; Camilli *et al.,* 2010; Shonkoff, 2010; McCoy *et al.,* 2017; Reynolds & Temple, 2019) as well as for taxpayers and society-at-large (Reynolds *et al.,* 2002; Belfield *et al.,* 2006; Heckman *et al.,* 2010; Cannon *et al.,* 2017, 2018; van Huizen *et al.,* 2019). However, despite this research, the level of public investments in early childhood education has been relatively low, leaving parents to shoulder the heaviest financial burden in a phase of life when they are less able to afford it (Council of Economic Advisers, 2016).

Some states in the U.S. fund preschool, but on average, among all 4-year-olds across the USA, before the pandemic only a third were enrolled in state preschool programs, and the states spend only a little over $5000 per child in preschool as compared to over $12,000 per year in elementary and secondary school (US Census Bureau, 2017; Friedman-Krauss *et al.,* 2019). Good quality early care and education are expensive, with the average full-time childcare costs exceeding the average cost of in-state college tuition in the USA (Schulte & Durana, 2016). Thus, conducting rigorous benefit-cost analysis to assess the magnitude of benefits relative to costs of early childhood programs can be of great importance to ongoing discussions of how many public resources to devote to these programs.

While a few notable studies have investigated impacts on health outcomes in adolescence and adulthood, the majority of the cost-benefit research so far has focused on outcomes such as educational attainment, earnings, and involvement in criminal activity (Reynolds *et al.,* 2002, 2011*b*; Heckman *et al.,* 2010). The economic impact of early childhood education on health outcomes in adulthood remains understudied, despite the adverse effects of poor health on state and federal budgets and the quality of life. Previous studies of the Chicago Child-Parent Center (CPC) program have identified pathways through which early education programs can impact the health-related outcomes through adulthood (Topitzes *et al.,* 2009; Reynolds & Ou, 2016). To the extent that future health benefits of early childhood programs are overlooked in economic evaluations, the total benefits to society are undercounted.

Through this study, we wish to examine the following research questions:

  (i) Is there a sustained impact of the Chicago CPC preschool program on physical and mental health outcomes through the mid-thirties?
 (ii) How large are the monetized benefits of improved health relative to program costs?

In our cost-benefit analyses, we follow the standards expected for conducting and reporting economic evaluations for preventive interventions (Crowley *et al.,* 2018) and specifically early childhood interventions (Karoly, 2012). We estimate the impact of the program on each outcome and use recommended methods in statistical inference for multiple comparisons to adjust *p*-values. We carefully define the health outcomes of interest to avoid the double-counting of benefits and then perform sensitivity analysis including Monte Carlo analysis and use a range of discount rates to test the robustness of our results. Moreover, our analysis differs from prior CPC work in two major ways. First, this report only examines the preschool component of CPC while the two prior studies also assessed school-age services from first to third grade and the extended (preschool to third grade) program for 4–6 years versus fewer years. In the present study, however, the estimates for preschool impacts are adjusted for participation in the school-age program for which both groups were eligible. Analyzing all program contrasts was beyond the scope of the present study. Second, we only examine impacts for physical and mental health outcomes, not all domains of well-being as the prior studies did. Such a study, which is now underway, would have greatly increased the length and complexity of the report.

2 Literature review

One reason for the underinvestment in good quality early education programs may be that while many studies report strong effects, the evidence is not entirely positive. Some well--publicized research shows that the early gains especially in terms of test scores that are evident in kindergarten may not be discernible by second or third grade. Two notable examples employing random assignment include the U.S. Department of Health and Human Services (HHS)'s Head Start Impact Study and the Tennessee Pre-Kindergarten (VPK) Program. In the Head Start Impact Study, families were randomly assigned an offer for enrollment in Head Start, a federally funded preschool program for children in low-income families. While the first report from the study in 2005 found small but improved effects on cognitive skills, the subsequent report in 2010 found that the effects had mostly faded out by the end of first grade and there was no impact on test scores by the end of third grade (Puma *et al.,* 2005, 2010, 2012; Bauer, 2019). Subsequent researchers have reanalyzed these data to incorporate significant heterogeneity in the counterfactual care settings available to study participants and found sizeable effects of Head Start participation for children who would otherwise have not attended preschool (Feller, Grindal, Miratrix, & Page, 2016; Kline & Walters, 2016).

In the evaluation of the Tennessee Pre-Kindergarten (VPK) Program, Lipsey *et al.* (2018) found that while the program had what Kraft (2020) describes as a large impact (effect size >0.25) on achievement measures in pre-k, none of these effects of the program on children's academic achievement and behavior persisted through the third grade. Some evidence indicates that the VPK program participants ended up performing worse than the comparison group on Mathematics (effect size: −0.23) and Science (ES: −0.20) scores in third grade. The authors did note that the quality score of the VPK program (4.15) was below the average of other state pre-k programs (4.35) when measured using the *Early Childhood Environmental Rating System-Revised (ECERS-R)*. Subsequent research by other researchers found that the VPK participants who had attended high-quality elementary schools with high-quality teachers had significant gains in math and language lasting through third grade (Pearman *et al.,* 2020).

While there have been numerous studies on the benefits of early intervention, only a subset of these studies has looked at the long-term impacts of early childhood education in later adulthood, well beyond high-school graduation. In this article, we limit our focus to studies that have evaluated the impact of an early childhood education program beyond 20 years. Further, we review the studies that investigate the long-term impact on health outcomes or have conducted cost-benefit analyses that may or may not include benefits accruing to improved health outcomes.

2.1 Long-term evaluations of major early childhood programs on non-health outcomes

The long-term effects of various early childhood education programs have been evaluated by multiple researchers using a variety of research designs. Estimates of the long-term effects of the Head Start program, which is a nationally offered federally funded preschool program for poor children in the US, tend to come from studies that leverage historical regional differences in program availability to estimate effects by comparing participants in national surveys who were more or less likely to have participated. Carneiro & Ginja (2014), Barr & Gibbs (2019), and Bailey *et al.* (2020) are examples of studies that find significant effects of Head Start on educational attainment and crime. Additionally, sibling comparison studies have found that children who attended Head Start were significantly more likely to complete high school by approximately 9 percentage points (Garces *et al.,* 2002; Deming, 2009), and

attend college by 6–12 percentage points (Deming, 2009; Bauer & Schanzenbach, 2016). These findings, however, should be interpreted with caution since these studies rely on retrospective reports of participation from data sets whose original purpose was not to assess the effects of Head Start. Prospective longitudinal cohorts studies of early childhood programs provide more interpretable estimates of impact given that participation and implementation is measured fully in real time.

More direct evidence of other early childhood programs comes from several longitudinal studies employing randomized trials or strong non-experimental study designs that follow participants into adulthood. Three early childhood education programs have been extensively evaluated by following a cohort of program and comparison group children over more than three decades – the HighScope Perry Preschool Program[1] (PPP), the Abecedarian Project[2] (ABC), and the Child Parent-Center Program[3] (CPC). Although the three programs differ in their scale of operation, location and time of the offering, target population, and programming, evidence of improved outcomes in later-life associated with early education intervention are found in studies of each of these programs, including but not limited to higher high school graduation rates, adult employment, earnings, and lower involvement in criminal activities (Campbell *et al.,* 2012; Schweinhart, 2013; Reynolds *et al.,* 2018; Heckman & Karapakula, 2019; García *et al.,* 2020).

2.2 *Study of health outcomes of early childhood programs*

Evaluations of the Head Start program have found the program to be associated with increased access to preventative health services (Currie & Thomas, 1995), improved measure of self-reported health status (Deming, 2009), lower incidence of depression (Bauer & Schanzenbach, 2016), lower incidence of behavioral problems, and reduction in obesity among adolescents (Carneiro & Ginja, 2014). Englund *et al.* (2014) investigated the effects of early childhood education on adult health (age 21) using data from the longitudinal studies of the PPP, ABC, and CPC programs. They measured outcomes such as alcohol, tobacco, and drug use, two or more indicators of these health-compromising behaviors, and depressive symptoms. The study found significant differences in outcomes for drug use (ABC, CPC), tobacco use (CPC), depressive symptoms (CPC), and on two or more indicators of health-compromising behaviors (CPC, PPP).

In the age-37 follow-up study of PPP, Muennig *et al.* (2009) found statistically significant improvements in the health status of program participants. However, the observed

---

[1] The Perry preschool program was conducted from 1962 to 1967 at the Perry Elementary School in Ypsilanti, Michigan. It was a 2-year program spanning 35 weeks in a year, offered for 2 and a half hours for 5 days a week for Black children considered to be at risk of school failure due to poverty. The study involved random assignment of 123 participants into the program (58 children), while the control group (65 children) was not offered any alternative programming (Schweinhart & Weikart, 1997).

[2] The Carolina Abecedarian Project (ABC) was offered from 1972 to 1977 in Chapel Hill, North Carolina. It was a 5-year program spanning 50 weeks in a year, offered for 8 hours a day for 5 days a week. The project randomized 111 infants into the program group (57 children) that received the comprehensive program, or the control group (54 children) that received only nutritional supplements and parental counseling (Campbell *et al.,* 2002).

[3] The Chicago Longitudinal Study (CLS) tracks a cohort of 1539 children who attended early childhood programs in Chicago, Illinois in 1983–1984. The intervention group (989 children) received the Child–Parent Center program for 35 weeks for 2 years, 2 and a half hours a day for 5 days a week, with school-age services during first to third grades. An appropriate comparison group (550 children) was created by matching on the age of kindergarten entry and demographic characteristics. About a fourth of the comparison group was enrolled in Head Start (Reynolds *et al.,* 2001).

differences in rates of self-reported medical conditions as diabetes, hypertension, obesity, asthma, and arthritis were either not significantly different or in the opposite direction of that expected. In a further follow-up at age 55, Heckman & Karapakula (2019) found health effects of the program in form of lower excessive cholesterol and lesser likelihood of being bedridden for males, and lower rates of diabetes and substance abuse among the females, as compared to the control group.

Muennig *et al.* (2011), in their age-21 follow-up study of the Abecedarian program, found that the program participants had improved health outcomes (consisting of depression index, health problems, and hospitalizations), and reduced chances of behavioral risk factors. Interestingly, a later study at age 30 by Campbell *et al.* (2012) found no significant effects of the program on mental health, substance abuse, and self-reported health status. However, the study could be limited by its small sample size to have enough power to detect small effects of the program which could also be significant. In a further follow-up study with a specific focus on health outcomes in mid-thirties, using medical exam data on blood pressure, hypertension, BMI, and cholesterol, Campbell *et al.* (2014) found that the program was associated with a significantly lower prevalence of health risk factors. In the most recent reanalysis of the study data, García *et al.* (2018) found that the ABC/CARE (Carolina Approach to Responsive Education) program had a significant effect on a composite health index of both male and female participants. The index included health outcomes and health behaviors including smoking, body mass index (BMI), psychological distress, asthma, high blood pressure, heart disease, cancer, lung disease, and diabetes, among others, at age 30. While the effect was statistically significant, it was not practically large with an effect size of 0.06.

Dietrichson *et al.* (2018) conducted a meta-analysis of 26 studies that evaluated long-term outcomes of universal early childhood programs across the world and found mixed effects on long-term health outcomes. However, a detailed reading of the study revealed that only three of the 26 studies included in the analysis actually reported measures of health and well-being. Nonetheless, these studies contribute to a small set of literature that evaluates a variety of health outcomes much later into adulthood.

2.3 Benefit-cost analyses of early childhood programs into adulthood

Ramon *et al.* (2018) conducted an economic review of 11 studies that evaluated early childhood education programs in the US across the district, state, and federal levels, and found median returns of $4.19 in total benefits for every dollar invested in a program. While the estimated benefits did not include health measures, the study argued from a health equity perspective that the benefits from a more productive labor force could provide cost savings for government health care programs and private health insurers.

Using data from the Head Start Impact Study (HSIS), Kline & Walters (2016) estimated a benefit-cost ratio of the Head Start program to be in the range of 1.10–1.84 including only the effects on adult earnings. It is worth noting that the benefits of the program could be much higher if the impacts on crime, health, or other externalities were also incorporated into the analysis.

Two major cost-benefit analyses of the Perry program were conducted at age 27 and age 40, where, the program was estimated to provide a return of $7.16 and $12.90 for every dollar invested in the program in the age-27 and age-40 follow-up respectively (Schweinhart, 1993; Barnett, 1996; Nores *et al.,* 2005). While the study accounted for

benefits accruing from the reduction in schooling and welfare costs, reduction in the justice system and victim costs from the crime reduction, and increase in taxes paid resulting from higher earnings, health benefits were not included in the analysis. Belfield *et al.* (2006) estimated the benefit-cost ratio going further higher to 12.90 in the age-40 follow-up. However, most benefits were accrued from lower criminal justice system expenditures, higher tax revenues, and lower welfare payments. While the study measured some health outcomes such as smoking, frequency of drug use, and drinking, these were not included in the benefit calculation due to a lack of significant impacts. A reanalysis of the benefits of the program by Heckman *et al.* (2010) produced a similar benefit-cost ratio of 12.20.

Barnett & Masse (2007) conducted a benefit-cost analysis of the ABC program at age 21 and found the program to repay \$2.5 for every \$1 invested. While they included the benefits from cost savings related to smoking, other health measures were not evaluated in this study. In the most recent follow-up of this sample, García *et al.* (2020) estimated the full life-cycle benefits and costs of the ABC/CARE (Carolina Approach to Responsive Education) programs by forecasting its full array of benefits such as labor income, parental income, health, and reduced crime. They estimate a benefit/cost ratio of 7.3 after adjusting for the welfare cost of taxation to society to fund the programs.

In benefit-cost analyses of the CPC program at ages 21 and 26, Reynolds *et al.* (2002, 2011) found the program to provide returns in the range of \$7–11 for every dollar invested. However, these studies did not measure the effect of the program on health outcomes, with the majority of the benefits accruing from increased earnings and tax revenues, and averted criminal justice system costs. While the effect on smoking was estimated in the age 26 study, the benefits were not included in the analysis as a statistically significant impact on smoking was not found. A partial CBA comparing the preschool program costs to the benefit solely arising from reduced expenditures on special education has been reported in Temple & Reynolds (2015). The study reports that the benefits in terms of special education cost savings amount to approximately 60 % of the preschool costs.

3 CPC program and pathways to long-term impact

The CPC program is a center-based early childhood intervention that provides comprehensive, continuous educational and family support services to economically disadvantaged children from preschool through third grade (Figure 1). As noted earlier, due to space limitations and clarity in reporting, we focus on the preschool component with adjustment for differential participation in the school-age program. The CPC program focused on five key features: providing an early education intervention no later than age 4, a structured learning approach for language and basic skills, increased parent involvement, provision of health and social services, and program continuity between preschool and elementary school (first to third grades) (Reynolds *et al.,* 2003). See also Reynolds & Ou (2016) for a historical account and manual for current implementation.

The program operated 3 hours daily for 5 days a week. To ensure individual attention for each child, the centers had a low child-to-staff ratio of 17:2. There was a special focus within the program on increasing the involvement of parents in the school, requiring them to participate at least one-half day per week in activities such as going on field trips or trips to the library, and supervising reading and play sessions at the school. The promotion of health and good nutrition was a component of the CPC program, with physical health and psychological development being one of the important goals of the program during its
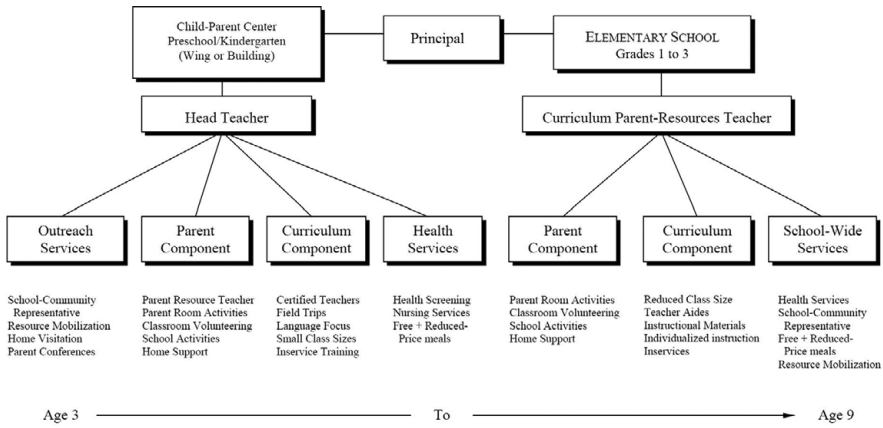
Figure 1. *Child-parent center program model. Reprinted from "Success in early intervention: The Chicago child-parent centers" by Reynolds (2000, p. 33). Copyright 2000 by the University of Nebraska Press.*

inception (Reynolds, 2000, p. 27). After enrolling in the CPC program, children were provided a health screening from a registered nurse on-site and given tests for vision and hearing. Medical and immunization histories of children were obtained from parents and they were required to have a physical and dental examination. Children who needed preventative services were then referred to appropriate service agencies (Reynolds, 2000, p. 43). The program also made available special medical and educational services such as speech therapists and school psychologists to the participants (Reynolds, 2000, p. 47).

To promote understanding of how early intervention can affect later outcomes, Reynolds (2000) developed the 5-Hypothesis Model (5HM) (Figure 2), where the promotion of cognitive and scholastic advantages, motivational advantages, social adjustment, family support behaviors, and school supports accounted for a majority of the estimated direct effect of preschool on adult well-being (Reynolds & Ou, 2016). The school support hypothesis was recently broadened to the school and community support hypothesis (Reynolds *et al.,* 2019). These pathways suggest the mechanisms through which the effects of an early childhood program are sustained into adulthood. In another study of the pathways of effects of the CPC program, Reynolds *et al.* (2004) found that higher attendance in high-quality elementary schools, lower mobility, increased literacy skills in kindergarten, avoidance of grade retention, and increased parent involvement in school as a result of CPC participation, mediated the long term effects on educational attainment and reduction in crime. Topitzes *et al.* (2009) conducted an exploratory path analysis of the CPC program to identify the effects of several education-related experiences on health-related outcomes through age 24 and found that the benefits of health-related outcomes in early adulthood can be attributed back to participation in a high-quality, comprehensive preschool program. They found that the program participants were 22 % less likely to smoke tobacco daily and were 20 % more likely to carry health insurance.

Thus, the focus on physical and mental health during the preschool program and the habits inculcated since early childhood days may have long-term effects on health outcomes in adults in addition to the impact on educational outcomes. In this study, we evaluate whether the CPC program impacts the adult health outcomes in mid-thirties and whether the benefits accrued from these improved health outcomes outweigh the program costs.
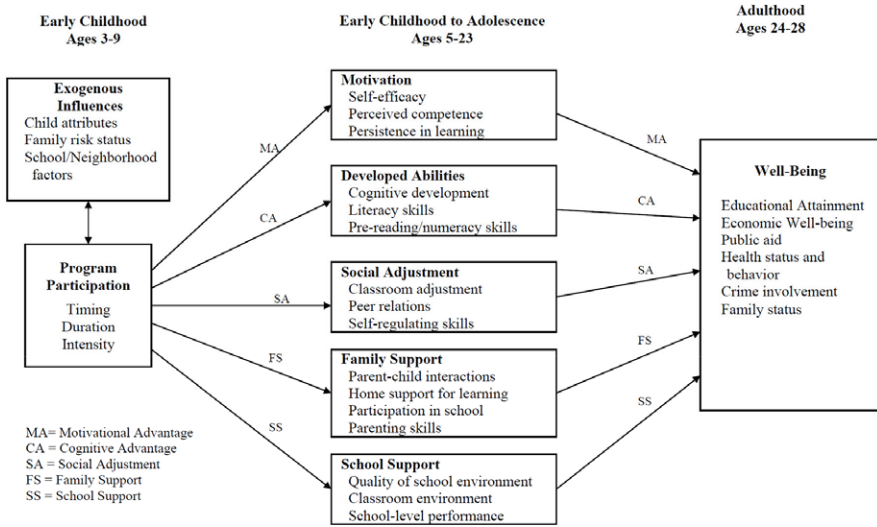
Figure 2. *Five-hypothesis model depicting the path from early childhood intervention to adult well-being. Reprinted from "Generative Mechanisms in Early Childhood Interventions: A Confirmatory Research Framework for Prevention" by Reynolds & Ou (2016, p. 798). Copyright 2015 by the Society for Prevention Research.*

4 Data

This study makes use of data from the Chicago Longitudinal Study (CLS), which follows a same-age cohort born in 1979–1980 and surveyed most recently at age 35–37. All participants in the CLS attended early childhood programs in public schools located in high-poverty neighborhoods in Chicago. The quasi-experimental study design was created by including the complete cohort of 989 children who enrolled in the federally-funded Title I CPCs beginning in at ages 3 and 4 (1983–1985). The remaining one-third of the study sample of 550 were children of the same age who attended the usual early childhood programs in the district, primarily full-day kindergarten in randomly selected schools. These schools participated in the Chicago Effective Schools Project (CESP), the existing government-funded alternative intervention designed to improve student achievement. Full-day kindergarten was the common early childhood intervention for economically disadvantaged students at that time. Nearly 15 % of the comparison group also attended Head Start preschool. Thus, the comparison group in this study received alternate intervention rather than no intervention and matched the poverty characteristics of the CPCs (Reynolds, 2000; Reynolds *et al.*, 2011*a*). CPC participants could continue their participation up to third grade for a total of six years if they began at age 3. Any study participant who enrolled or moved to a CPC school received services. Our focus in this study is CPC preschool and we controlled for later school-age participation, of which 55% of the study sample enrolled for at least one year (n = 850).

Considering the non-randomized assignment of the children to the program or comparison group, there could be a possibility of selection bias challenging the group comparability. We use a rich array of data on demographic characteristics as statistical control for self-selection on observables and reweight the individual observations to achieve covariate balance in both treatment assignment and attrition as in Reynolds *et al.* (2011*a*).

Table 1. *Baseline characteristics of the study sample by the program status.*[a]

Characteristic	Total (*n* = 1539)	CPC (*n* = 989)	Comparison (*n* = 550)
Female	50.1	51.8*	47.1
African American	93.0	92.7	93.5
Mother not completed high school by child's age 3[b]	54.3	51.0***	60.2
Ever reported receiving free lunch by child's age 3[b]	83.8	84.2	82.9
Mother under 18 at child's birth[b]	16.1	15.6	17.3
Having four or more children at home by child's age 3[b]	16.6	16.0	17.8
Ever reported receiving AFDC by child's age 3[b]	62.8	63.1	62.2
Mother not employed by child's age 3[b]	66.3	67.3	64.6
Single parent by child's age 3[b]	76.5	76.7	76.0
Reside in high poverty school area[b,c]	76.0	77.7**	72.9
Missing on any family risk indicators	16.2	14.8**	18.9
Child low birth weight (<2500 g)	11.8	10.9	13.3
Family conflict, child age 0–5	5.7	5.7	5.9
Family financial problems, child age 0–5	7.0	7.1	6.8
Substance abuse parent, child age 0–5	4.1	4.3	3.9
Full-day kindergarten	74.1	59.8***	100.0
CPC school-age intervention (1–3 years)	55.2	69.2***	30.2
CPC extended intervention (4–6 years)	35.9	55.9***	0.0

Significance level: ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

[a]Data are presented as a percentage of individuals. Administrative records (birth certificates and school records) were used for information on the child and family background indicators from birth to 5 years of age. Home environment problems and adverse childhood experiences were obtained from retrospective reports.

[b]Family risk indicators.

[c]Children in school areas in which more than 60 % of children reside in low-income families.

A summary of the characteristics of the CPC and comparison groups is shown in Table 1. The study participants were predominantly African American (93 %) and resided in the most disadvantaged neighborhoods of Chicago that had nearly double the neighborhood poverty rate as compared to the city average. Children in the program group and comparison group were equivalent on child and family characteristics such as gender, single-parent family status, enrollment as 3-year-olds, and receipt of special education services. We account for the difference in some of the child and family risk characteristics between the intervention and the control group by including these as covariates in the main analysis, including CPC school-age program participation.

The data were collected through numerous sources including but not limited to birth and school records; child maltreatment and justice system records; educational attainment, adult earnings, employment, and public aid; and participant and family surveys. Data collected from these sources has been used in the past research of the CPC program (Reynolds *et al.*, 2011*a*, 2018). The data on health outcomes were collected as part of a comprehensive interview during the follow-up at age 35–37 which included 130 questions about physical

and mental health, economic well-being, life history, and more. The survey was conducted between 2012 and 2017 and over 80 % of the respondents completed the interview via telephone. Ou *et al.* (2020) provide a detailed description of the approach that was used to locate and interview the participants in this study.

Follow-up data at age 35 were available for a total of 1125 participants (retention rate = 73 %), 741 across the program group (75 % retention), and 384 across the comparison group (70 % retention). As reported in Ou *et al.* (2020), the retention rate in our study is higher than other large-sample (>200 participants) early childhood intervention studies such as the Infant Health and Development Program (65 % up to age 18), the Houston Parent-Child Development Center (63 % up to age 18), the Early Head Start Research and Evaluation Project (54 % up to age 10), and the Consortium for Longitudinal Study (55 % up to age 22).

Table 2 compares the child and family characteristics of the in-sample and attrition groups. Non-respondents were more likely to be males, were born into a family that participated in the AFDC program, had mothers who did not complete high school or were teenage at birth, and did not participate in the CPC program. The right-hand side of the table shows the results of using the Inverse Propensity Score Weighting (IPW) procedure, described in the next section. Table 3 reports the covariate means for those students who attended the CPC program in preschool and those who did not. A few differences are apparent, specifically for sex, maternal education, and residence in a very high-poverty neighborhood. In Tables 2 and 3, calculating the means after the reweighting of observations via the IPW procedure results in close covariate balance as might be expected if treatment assignment and attrition were random.

The use of self-reported measures of health outcomes is common in early childhood research as conducting medical exams for all participants could be extremely challenging and expensive. Muennig *et al.* (2009) used self-reported data on health conditions including diabetes, hypertension, obesity, asthma, and arthritis in their age-37 follow-up study of Perry Preschool. Similarly, Campbell *et al.* (2012) used self-reported health status and measures of drug, alcohol, and tobacco consumption in their age-30 follow-up study of the Abecedarian program. Although in a later study, Campbell *et al.* (2014) were able to use clinical measures of outcomes such as blood pressure, hypertension, BMI, and cholesterol, the study had a limitation of a lower response rate (~61 %). Additionally, in our study, the measure of BMI created using self- reported height and weight data obtained from the survey was highly correlated ($r = 0.85$) with an in-person BMI measurement conducted with a participant subsample (Eales *et al.*, 2020). We consider the self-reported data used in our study to be a reliable source for health outcomes data and for a subsample of participants these measures can be compared to outcomes obtained through in-person exam data.

4.1 Outcome measures

We analyze the physical and mental health outcomes for participants interviewed as of age 37. First, we created a dichotomous variable for smoking which indicates whether the participants currently smoke any tobacco product more than once a day. Similarly, we created an indicator of drug use based on whether they have ever used drugs harder than marijuana. Next, we used the self-reported information on height and weight to calculate the body mass index (BMI), which is then used to create an indicator of obesity (obesity = 1, if BMI > 30). Finally, we create an indicator for each of diabetes, hypertension, and depression

Table 2. *Means of covariates without and after inverse probability weighting for attrition.*

25 Predictors used for Attrition[a]	Unweighted				Weighted			
	Not in Sample[b]	In Sample	Difference	p-value	Not in Sample	In Sample	Difference	p-value
Mother not completed high school, child age 0-3, %	58.8	52.4	6.4	0.022**	53.1	53.9	−0.8	0.790
Child eligible for subsidized meals, child age 0-3, %	85.7	83.0	2.7	0.179	82.3	83.5	−1.2	0.627
Mother under age 18 at childbirth, %	18.8	15.1	3.7	0.079*	16.1	16.2	−0.1	0.942
Four or more children in the family, child age 0-3, %	16.5	16.7	−0.2	0.937	17.1	16.8	0.3	0.889
Participate in AFDC program, child age 0-3, %	66.7	61.2	5.5	0.041**	61.5	62.4	−0.9	0.781
Mother not employed, child age 0-3, %	68.1	65.6	2.5	0.355	64.6	65.9	−1.3	0.678
Single parent family status, child age 0-3, %	78.7	75.6	3.1	0.186	75.3	76.2	−0.9	0.730
Indicator for missing risk factors, child age 0-3, %	20.6	14.5	6.1	0.003***	15.3	15.6	−0.3	0.876
Reside in high poverty school area, %	77.1	75.5	1.6	0.488	76.2	76.1	0.1	0.968
Low birth weight (<2500g), %	9.7	12.6	−2.9	0.116	11.9	11.9	0.0	0.997
Family conflict, child age 0-5, %	4.6	6.2	−1.6	0.229	6.5	5.8	0.7	0.689
Family financial problems, child age 0-5, %	5.8	7.5	−1.7	0.236	7.4	7.1	0.3	0.826
Substance abuse parent, child age 0-5, %	3.5	4.4	−0.9	0.415	4.3	4.2	0.1	0.906
Female child, %	39.6	54.3	−14.7	0.000***	50.6	50.3	0.3	0.897

Table 2. *Continued*

25 Predictors used for Attrition[a]	Unweighted				Weighted			
	Not in Sample[b]	In Sample	Difference	p-value	Not in Sample	In Sample	Difference	p-value
African American child, %	91.2	93.7	−2.5	0.078*	93.1	93.0	0.1	0.937
CPC preschool program participation, %	60.0	66.0	−6.0	0.025**	64.0	64.3	−0.3	0.912
CPC School-age program participation, %	50.9	57.0	−6.1	0.030**	55.6	55.5	0.1	0.985
Standardized word test, child age 5	61.6	64.7	−3.1	0.000***	64.2	63.9	0.3	0.674
Proxy of residential mobility	1.0	0.9	0.1	0.005***	0.9	0.9	0.0	0.864
Census tract neighborhood mobility < 1 year, %	19.1	18.8	0.3	0.497	18.8	18.8	0.0	0.846
Census tract neighborhood mobility 1-5 years, %	29.4	29.3	0.1	0.775	29.3	29.3	0.0	0.991
Census tract neighborhood mobility 5-10 years, %	22.8	22.9	−0.1	0.807	23.2	23.0	−0.2	0.684
Census tract neighborhood mobility 10-20 years, %	24.8	25.3	−0.5	0.399	25.0	25.2	−0.2	0.869
Census tract self-employed rate, %	1.9	1.8	0.1	0.201	1.8	1.8	0.0	0.817
Census tract African American female householder, %	40.3	39.5	0.8	0.356	39.8	39.8	0.0	0.985

Significance Level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

[a]Data are presented as a percentage of individuals except for standardized word test and proxy of residential mobility.

[b]Health outcomes data for 414 participants out of the original sample of 1539 were not available. In-sample group consisted of 1125 participants (program: 741, comparison: 384).

Table 3. *Means of covariates without and after inverse probability weighting for program selection*

Predictors for Program Selection[a]	Unweighted				Weighted			
	CPC[b]	Comparison	Difference	p-value	CPC	Comparison	Difference	p-value
Females, %	51.8	47.1	4.7	0.080*	49.9	49.5	0.4	0.878
African American, %	92.7	93.5	−0.8	0.589	93.1	93.4	−0.3	0.822
Mother not completed high school by child's age 3, %	51.0	60.2	−9.2	0.001***	54.0	54.0	0.0	0.998
Ever reported receiving free lunch by child's age 3, %	84.2	82.9	1.3	0.502	83.6	83.6	0.0	0.995
Mother under 18 at child's birth, %	15.6	17.3	−1.7	0.385	16.3	16.2	0.1	0.964
Having 4 or more children at home by child's age, %	16.0	17.8	−1.8	0.353	16.9	16.8	0.1	0.946
Ever reported receiving AFDC by child's age 3, %	63.1	62.2	0.9	0.723	62.6	62.8	−0.2	0.948
Mother not employed by child's age 3, %	67.3	64.6	2.7	0.266	66.1	66.1	0.0	0.998
Single parent by child's age 3, %	76.7	76.0	0.7	0.742	76.5	76.6	−0.1	0.964
Missing on any family risk indicators, %	14.8	18.9	−4.1	0.035**	15.6	15.4	0.2	0.916
Reside in high poverty school area, %	77.7	72.9	4.8	0.037**	76.2	76.3	−0.1	0.954
Child low birth weight (<2500g), %	10.9	13.3	−2.4	0.170	11.7	11.7	0.0	0.976
Family conflict, child age 0-5, %	5.7	5.9	−0.2	0.856	5.7	5.7	0.0	0.986
Family financial problems, child age 0-5, %	7.1	6.8	0.3	0.842	6.9	6.8	0.1	0.922
Substance abuse parent, child age 0-5, %	4.3	3.9	0.4	0.718	4.1	4.0	0.1	0.950

Significance Level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
[a]Data are presented as a percentage of individuals.
[b]Out of 1539 study participants, 989 participants were in the CPC group and 550 participants in the comparison group.

Table 4. *Unadjusted Mean Comparison of Health Outcomes*

Variable Description[a]	(1) Total Sample	(2) CPC Sample[b]	(3) Comparison Sample[c]	(4) Difference in Mean (Standard Error)
Smoking[d]	21.5 $n = 1100$	19.7 $n = 722$	24.9 $n = 378$	−5.2** (0.026)
Drug Use[e]	5.8 $n = 1097$	5.0 $n = 721$	7.5 $n = 376$	−2.5* (0.015)
Body Mass Index[f]	30.49 $n = 1065$	30.29 $n = 704$	30.86 $n = 361$	−0.57 (0.446)
Obesity (BMI > 30)	45.4 $n = 1065$	44.2 $n = 704$	47.6 $n = 361$	−3.4 (0.032)
Diabetes[g]	5.4 $n = 1097$	4.1 $n = 724$	7.8 $n = 373$	−3.7** (0.014)
Hypertension[g]	16.9 $n = 1096$	16.7 $n = 723$	17.1 $n = 373$	−0.4 (0.024)
Depression[g]	12.5 $n = 1098$	13.0 $n = 723$	11.5 $n = 375$	1.5 (0.021)

Significance Level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
[a]Data are presented as percentage of individuals except for Body Mass Index.
[b]CPC = Child-Parent Center program.
[c]The comparison group was created from a matched set of similar high-poverty schools.
[d]The response was coded as 1 if currently smoke any tobacco product more than once a day.
[e]The response was coded as 1 if used drugs harder than Marijuana.
[f]BMI measure was created using self-reported measures of height and weight.
[g]The response was coded as 1 if the participant was ever diagnosed with the condition.

using the self-reported survey responses of whether the respondent was ever diagnosed with these conditions.

Table 4 provides information on the outcomes examined and how their means vary across the CPC participants and the comparison group. As shown in Table 4, there are significant differences with respect to smoking, drug use, and diabetes between the CPC and comparison groups.

5 Method

Consistent with impact evaluations previously conducted with this sample (Reynolds *et al.*, 2011*a*, 2018) we used probit and linear regression to estimate the marginal effects of CPC participation on the health outcomes. The covariates include child and family characteristics such as race/ethnicity, gender, childbirth weight, receipt of child welfare services, parent education, single-parent family status, teen parenthood, employment, four or more children in the family, and school-poverty rate of the kindergarten sites, among others. With the exception of public aid receipt and eligibility for subsidized meals, these control variables were measured from birth to age 3. To estimate the effect of the different components of the preschool and school-age intervention through third grade, indicators for participation in these components of the PK-3 intervention were included in the regressions.

A key threat to internal validity in any longitudinal study is non-random attrition, where the concern is that participants who leave the sample can have different observed and unobserved characteristics than those who remain, and non-random assignment, where the program participants have different characteristics than the comparison group, which can affect the program estimates. To avoid these concerns of non-random program assignment and attrition, we use Inverse Probability Weighting (IPW), a method of weighting the observations by the inverse probability of each observation being missing conditional on covariates as a regression weight for both attrition and program assignment. Impact evaluations using observational data are increasingly incorporating propensity score methods for estimation of program effects (Austin & Stuart, 2015). An advantage of weighting over matching is that the latter often discards observations and researchers have suggested that IPW methods may have efficiency advantages (Hirano *et al.*, 2003) as long as the trimming of extreme weights is undertaken (Austin & Stuart, 2015).

The IPW results make use of the larger analytic sample to better understand and control for the differences in characteristics of students for whom we have data on health outcomes compared to those who are missing this data (Seaman & White, 2013). To do this, we first run a probit regression using the larger sample of students with and without health outcomes data at age 37 to generate predicted probabilities for each student of having this information using observable characteristics assumed to influence attrition. Being able to predict the missingness of the dependent variable is important. Hence the prediction model includes a comprehensive set of child and family risk variables that also will be included in the main outcomes regressions as well as a set of additional variables that are listed in Table 10 of the Appendix. With the results of the probit estimation, we then assign a weight of $1/p1$ to each child, where $p1$ is the predicted probability of the child being in the recovery sample ($R = 1$; otherwise 0). The first propensity score is estimated as $P1i = \Pr(R = 1/X)$. Similarly, we estimate a propensity score for participation in the program ($T = 1$; otherwise 0) as $P2i = \Pr(T = 1/X)$. As a double adjustment, the weights were multiplied together to produce a combined propensity score model as $P3i = P2i \times P1i$ following the previous research (Reynolds *et al.*, 2011a). Further, we used robust standard errors including the clustering of standard errors at the school level to account for within-site correlation in errors.

Students with characteristics associated with a lower predicted probability of being included in the analysis, but who actually have information for the health outcome, are rarer and consequently are given larger weight. Students with characteristics associated with a higher predicted probability of being in the regression are assigned smaller weights. In a sense, the use of IPW reweights the sample to better resemble a sample in which the outcome is missing at random. Since extreme probability scores (close to 0 or 1) producing large weights can yield unstable estimates, we set a floor and ceiling of weights by trimming them at 0.05 and 0.95 (Austin & Stuart, 2015).

The underlying idea behind the IPW approach is that the reweighting to remove systematic differences between the baseline mean observed characteristics of the program and comparison groups with a rich set of covariates creates a study design that more closely resembles a randomized trial (Austin & Stuart, 2015). While creating group equivalence for participants and controls (and for the retained participants versus those with missing outcomes) on covariates is the intended result, researchers may hope that the groups also are balanced on unobservable baseline traits, even though Austin *et al.* (2005) point out that this is not necessarily the case. We test for balancing of observed covariates in our

Table 5. *Effect of CPC on health outcomes after adjustments*

Health Outcome[a]	(1) Sample Mean (Sample Size)	(2) Unadjusted group difference (Standard Error)	(3) IPW adjusted Regression estimate (Robust SE)[b]
Smoking[c]	21.5	−5.2**	−5.8*
	n = 1100	(0.026)	(0.031)
Drug Use[d]	5.8	−2.5*	−2.5
	n = 1097	(0.015)	(0.018)
Body Mass Index[e]	30.49	−0.57	−0.96*
	n = 1065	(0.446)	(0.528)
Obesity (BMI > 30)	45.4	−3.4	−4.7
	n = 1065	(0.032)	(0.035)
Diabetes[f]	5.4	−3.7**	−4.1**
	n = 1097	(0.014)	(0.018)
Hypertension[f]	16.9	−0.4	−0.01
	n = 1096	(0.024)	(0.027)
Depression[f]	12.5	1.5	1.1
	n = 1098	(0.021)	(0.025)

Statistical Significance Levels *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
[a]Health outcome data was obtained from self-reported surveys between ages 35 to 37. Data are presented as the percentage of individuals except for Body Mass Index.
[b]Regression with covariates adjusted with Inverse Probability Weighting (IPW) for program selection and attrition. Standard errors clustered at the preschool site in parentheses.
[c]The response was coded as 1 if currently smoke any tobacco product more than once a day.
[d]The response was coded as 1 if used drugs harder than Marijuana.
[e]BMI measure was created using self-reported measures of height and weight.
[f]The response was coded as 1 if the participant was ever diagnosed with the condition

subsamples, as presented in the columns to the right side of Tables 2 and 3 and find no significant differences between the treatments and controls, and between the retained and dropped samples after weighting.

Once the estimates of CPC effectiveness are obtained in the regression analyses, the economic benefit is estimated by multiplying the marginal effect with the respective monetary estimate of the outcome. The present values of the monetized benefits are then compared to the costs of the intervention to compute the benefit-cost ratio, and the costs are subtracted from the benefits to calculate the net benefit.

6 Results

Results from our inverse propensity score weighted estimation approach (Table 5 and Table 11 of the Appendix) suggest a relationship between preschool participation and smoking, diabetes, and BMI. The CPC program group was 5.8 percentage points less likely to smoke daily as compared to the comparison group. The prevalence of diabetes was 4.1 percentage points higher among the participants in the comparison group as opposed to the CPC program participants. There was also a statistically significant association between

program participation and the body mass index. The differences in prevalence for diabetes were statistically significant at 5 % level, while the results on smoking and BMI were significant at 10 % level. No significant differences were found in drug use, the prevalence of obesity, hypertension, or depression.

In the economic evaluation that follows, we focus on the estimated effects of participation in the preschool program on reductions in the prevalence of diabetes and the rate of smoking. While the effect of program participation on BMI is marginally significant, we expect that the health consequences of having diabetes overlap in important ways with the consequences of being overweight. To avoid double-counting, we focus solely on reductions in smoking and diabetes diagnoses.

6.1 Controlling for multiple hypothesis testing

Since we test multiple non-independent health outcomes in our analysis, there could be a possibility of finding statistically significant results (false positives) for some of the outcomes merely by chance. We apply the Benjamini-Hochberg (BH) procedure (Benjamini & Hochberg, 1995) to control for multiple hypothesis testing as it has been used in other early childhood education studies (Knight *et al.*, 2019) and as Schochet (2008) points out, can have more statistical power than the Bonferroni method if any positive impacts of program participation truly exist. The BH procedure is easy to perform as it is based entirely on individual *p*-values. To perform the BH procedure, we first arrange the individual *p*-values associated with the results in the final column of Table 5 in ascending order and rank-order them from $i = 1$ to m (where i is the rank-order of the *p*-value from smallest to largest, and m is the total number of significance tests performed). Then we compare each individual *p*-value to the adjusted significance value, which is equal to $(i/m)Q$, where Q is the chosen false-discovery rate. The largest *p*-value for which $p < (i/m)Q$ is considered to be statistically significant, including all the *p*-values that are smaller than it.

The choice of the false-discovery rate Q is subjective and depends on the researchers' tolerance for the proportion of false positives, specifically the share of declared positives that are truly negative. Because prior studies based on RCTs of high-quality early childhood programs have reported positive health outcomes in adulthood (Muennig *et al.,* 2009, 2011; Campbell *et al.,* 2014), an expectation of 7 truly null hypotheses for the complete set of health outcomes seems overly conservative. As discussed above, even though 3 of the 7 outcomes were statistically significant in Table 5, the economic evaluation is only based on 2 of the 3 significant findings. Ex ante, we chose a false-discovery rate of $Q = 0.20$, which in essence means we are willing to allow 20 % of the three significant findings in Table 5 to be false positives. Table 6 shows the BH corrected critical *p*-value for each of the outcomes using a false-discovery rate $Q = 0.2$ and $m = 7$. We find that the program's effect on diabetes, smoking, and BMI is statistically significant even after using this correction for multiple hypothesis testing.

7 Benefit-cost analysis

Following the standard procedure employed in most benefit-cost analysis studies, the main steps in calculating costs and benefits of participation in the CPC program in terms of health

Table 6. *Benjamini-Hochberg corrected* p-*value.*

	(1)	(2)	(3)
Health outcome	*p*-value	Rank (*i*)	(*i/m*)*Q*
Diabetes	0.020	1	0.029
Smoking	0.063	2	0.057
Body mass index	0.069	3	0.086
Drug use	0.176	4	0.114
Obesity (BMI > 30)	0.183	5	0.143
Depression	0.656	6	0.171
Hypertension	0.996	7	0.200

Note: The table uses the value of $Q = 0.2$ (False Discovery Rate) and $m = 7$ (number of outcomes).

outcomes were as follows: (*a*) the costs and health benefits of the program are calculated in dollar terms, (*b*) to adjust for inflation, estimates are converted to 2021 dollars using the Bureau of Labor Statistics' Consumer Price Index for All Urban Consumers (CPI-U), (*c*) a discount rate of 3 % was applied to calculate the present value of the costs and benefits at age 3 (start of the program), and (*d*) the net present value of the program per participant was calculated by subtracting the present value of program costs from the present value of program benefits. Additionally, the program benefits were divided by costs to obtain the benefit-cost ratio (return for every 1 dollar invested). While previous studies (Barnett, 1996; Reynolds *et al.,* 2011*b*; Cannon *et al.,* 2018) focusing on adult earnings projected the benefits through age 65, we use the Centers for Disease Control and Prevention (2017) estimates which peg the average life expectancy of African-Americans born in the 1980s to be 68 years, since our sample consists of participants born in 1979 or 1980 who are primarily (93 %) Black. Further, we use 62 years as the age of retirement, as people can begin receiving their social security retirement benefits at that age. We used alternative discount rates (1, 5, and 7 %) as part of the sensitivity analysis.

7.1 Program costs

Two previous cost-benefit analyses have calculated the benefits and costs associated with the Chicago CPC program (Reynolds *et al.,* 2002, 2011*b*). The costs were primarily calculated using the operational budget of the program for the year 1985–1986 including the costs of instructional staff, staff for family and community support, administration, operations and maintenance, program materials, transportation, food, and community services, school-wide services, and school district support. Additionally, the cost of parent's time was estimated using the minimum wage for 10 hours of participation in the program per month. Imputed costs of capital depreciation and interest was also taken into account as suggested by Levin et al. (2017). Based on the cost estimates presented in these studies, the present value cost for 1 year of the CPC preschool intervention provided in 1986 was estimated to be $7233 per student in 2021 dollars.

All of the preschool participants received 1 year of preschool while 55 % received an additional year of intervention. Assuming the students that received only 1 year of preschool intervention entered the program at age 4, while those who received 2 years of the program

started at age 3, the per participant present value of the average cost of the preschool program evaluated at age 3 is (0.55)($7233) + (1.00)($7233/1.03) = $11,000.

7.2 Program benefits

The benefits from reduction in diabetes and smoking can be estimated using two methods: first, through adding the costs of medical expenditure, lost productivity, and the cost of premature mortality attributable to the condition, and second, through accounting for the economic costs of a reduction in quality of life in addition to the healthcare costs.

7.2.1 Benefits from reduced diabetes

Medical expenditure, lost productivity, and premature mortality cost of diabetes. We first estimate the benefits of reduced medical expenditures associated with a diabetes diagnosis using information from the recent study by the American Diabetes Association (2018). In 2017, the additional annual per-capita health care expenditures for people diagnosed with diabetes was $9601 or $10,342 in 2021 dollars. Because medical expenditures rise over time as diabetics age (Trogdon & Hylands, 2008), we adjust the estimates of the health care expenses to allow for this annual increase. Trogdon & Hylands (2008) found that after controlling for several medical conditions that develop over time with diabetes, annual health care costs associated with diabetes rise by $75 per year (or $93 in 2021 US dollars). Therefore, we adjust the cost of health care expenses by applying an age-based escalator of $93 per year to account for higher health care costs among older diabetics to estimate the life-cycle benefits from the reduction in diabetes.

The ADA report also includes estimates of the lost productivity due to a diabetes diagnosis. In 2017 dollars, they report an annual per-capita loss of $2830 associated with reduced productivity at work including the inability to work, or $3048 in 2021 dollars.

We make several assumptions in order to estimate the present value at age 3 of a future stream of improved productivity due to reductions in the probability of a diabetes diagnosis. As discussed above, we take the average life expectancy of our sample to be 68 years, and the retirement age as 62 years. Additionally, among our sample, the survey responses suggest that the average age of diabetes diagnosis was 28. We calculate the stream of benefits over the life course of the individual using the above estimates at a 3 % discount rate from age 28 to age 68 for health care costs, and up to age 62 for reduced productivity. These present value estimates of a lifetime of higher health care costs and reduced productivity associated with a diabetes diagnosis come out to be $136,207 and $32,218 respectively. For a 4.1 percentage point reduction in diabetes estimated to arise from participation in the CPC program, the present value of these benefits at age 3 is estimated to be $5584 and $1321 respectively.

The next step is to include the economic costs of early mortality due to diabetes. American Diabetes Association (2018) used the data from CDC National Vital Statistics Reports for total deaths to estimate the number of deaths primarily attributable to diabetes, and the proportion of deaths due to renal disease, cerebrovascular disease, and cardiovascular disease attributable to diabetes. They then estimated the present value of their foregone future earnings. While the average cost of premature mortality declines with age, American Diabetes Association (2018) estimated it to average $71,700, or $77,234 in 2021 dollars. Discounting this cost to the present value at age 3, and multiplying it with the program

impact of 4.1 %, we find the benefits from averted premature deaths to be $554 due solely to CPC participation.

Adding together the present values of reduced medical expenditures, improved productivity, and reduced premature mortality due to diabetes, we estimate the present value of the effect of the preschool program on the probability of being diagnosed with diabetes to be $7459.

Reduction in quality of life/utility value due to diabetes. A second way of monetizing the preschool impacts on diabetes is to consider the effect of diabetes on the quality of life in addition to the healthcare costs. Narayan *et al.* (2003) estimated the quality-adjusted life years (QALY) lost due to diabetes across a range of age of diagnosis, gender, and race. Their study highlighted significant racial disparities in health outcomes. For the non-Hispanic Black population diagnosed with diabetes at age 30, they estimated an average loss of 26.15 QALYs (24.2 for men and 28.1 for women) and an average loss of 18.5 total life years lost. While these estimates are almost two decades old, they are still relevant for our sample population, as in a more recent study, Rhodes *et al.* (2012) estimated 22.44 QALYs lost due to Type 2 diabetes in a cohort of 15–24-year-old adolescent Americans. Using the commonly used monetary value of $50,000 per QALY (Grosse, 2008), for a 4.1 percentage point reduction in diabetes attributable to the program, we estimate the benefits from the avoided loss of QALYs at age 30, discounted to the present value at age 3, to be $24,134.[4]

An alternate set of literature estimates the QALYs lost due to Diabetes using utility values obtained from using EuroQol five-dimensional (EQ-5D) valuation questionnaire. As opposed to an ideal utility value of 1 for perfectly healthy individuals, researchers have estimated the utility value for individuals diagnosed with diabetes in the range of 0.74 (Clarke *et al.,* 2002; Neumann *et al.,* 2014) to 0.80 (Huang *et al.,* 2007; Zhang *et al.,* 2012). In a meta-review that assessed 61 studies evaluating the utility values corresponding to diabetes complications, Beaudet *et al.* (2014) recommend a proposed utility value of 0.785. Using the utility value of 0.785 and a QALY estimate of $50,000 we estimate the life-cycle benefits from the aversion of reduced utility values between age 28 and 68 to be $5076 after discounting to the present value at age 3.

Adding the additional healthcare costs attributable to diabetes of $5584 as obtained above, the QALY approach yields a range of estimates of the economic benefits of a reduced rate of diabetes diagnoses due to preschool between $10,660 to $29,718 (or a mean of $19,935).

The discussion above suggests two approaches to monetizing the benefits of preschool participation in terms of its estimated impact on a diagnosis of diabetes by age 30 specifically tailored to this population of Black study participants born four decades ago. The first approach is to include averted costs from medical expenditure, reduced productivity, and premature mortality. The second approach is to include averted medical costs and improvements in QALYs. As explained by Shiroiwa *et al.* (2013), it is important not to add the improvement in terms of QALYs to the improved productivity estimates. Combining the two estimates would result in overcounting the benefits to the extent that the estimated utility value of 0.785 for diabetes includes concerns by the survey

---

[4] (26.15 QALYs $\times$ $50,000 $\times$ 4.1%)/(1.03^27).

respondents that this medical condition would reduce work productivity. For our cost-benefit analysis, we use the more conservative value of benefits from the reduction in diabetes to be $7459.

7.2.2 *Benefits from reduced smoking*

The approach to estimating the economic benefits associated with a reduction in smoking is similar to that used above for diabetes. We use the relevant results from a variety of studies that have estimated the economic benefits accruing from a reduction in the rates of smoking in adults in the form of savings in direct medical costs, lost productivity, and reduced mortality costs, and also consider the improvement in QALYs.

Medical costs, lost productivity, and premature mortality costs of smoking. Centers for Disease Control and Prevention (2008) estimated the average annual healthcare expenditure attributed to smoking to be $96 billion per year between 2000 and 2004. Another Centers for Disease Control and Prevention (2005) report estimated that approximately 44.5 million U.S. adults were smokers in 2004. Using the above two pieces of information,[5] we calculate the per capita annual incremental medical cost of smoking to be $2157, equivalent to $3047 in 2021 US dollars after adjusting for inflation.

Centers for Disease Control and Prevention (2008) also reports annual productivity losses from smoking-attributable diseases to be approximately $97 billion. Similar to the calculations above, we obtain the per-capita annual productivity losses attributable to smoking to be $2180 or $3079 in 2021 US dollars.

We use similar assumptions for estimating the cost of smoking as used in the section above while estimating the cost of diabetes. The average age around which Americans begin to smoke daily is estimated to be 18 years (U.S. Department of Health and Human Services, 2014). Additionally, studies have found that smokers lose at least 10 years of life expectancy as compared to non-smokers (Jha *et al.,* 2013; U.S. Department of Health and Human Services, 2014). Taking the normal life expectancy of our sample as 68 years as explained in the previous section, we assume the average life expectancy of smokers in our sample to be 58 years. Therefore, we estimate the stream of benefits over the life course of the individual due to averted medical costs and reduced productivity estimates attributable to smoking over the ages of 18–58, discounted to its present value at age 3. Using the discount rate of 3 %, the present value estimates of a lifetime of increased medical expenses and reduced productivity associated with smoking come out to be $47,163 and $47,658 respectively. For a 5.8 percentage point reduction in smoking-attributable to CPC program participation, the present value of these benefits at age 3 is estimated to be $2735 and $2764 respectively.

While the above estimates include medical costs and productivity losses, they do not include the cost of premature mortality attributable to smoking. We use two methods to estimate a range of benefits accrued from reduced mortality costs attributable to smoking. First, we estimate the foregone earnings due to premature death as the mortality cost. Based on our survey estimates, the average earnings for the preschool participants group at age 34 is $20,887 as compared to the average earnings of the comparison group which is $18,248 in

---

[5] $96 billion/44.5 million = $2157.

2017 US dollars. Among our response sample, 65.6 % of participants attended the CPC program, which brings the weighted average earnings of our sample to be $19,979 (equivalent to $21,521 in 2021 US dollars). We then estimate a stream of this earning up to the retirement age of 62 by allowing it to grow at 2 % per year. We assume that due to premature mortality, smokers will lose 10 years of life, however, they will lose only 5 years of earnings before retirement (between ages 58 to 62). The present value at age 3 of the last 5 years of earnings is estimated to be $33,401, which translates to a mortality-cost savings worth $1937 corresponding to a 5.8 percentage point reduction.

The alternate method to estimate the mortality costs is using the value of statistical life (VSL). VSL estimates vary depending on the characteristic of a disease, and due to a shortage of condition-specific VSL studies in the literature it is relatively uncommon to be able to use them in studies of cost of a particular illness (Peterson *et al.,* 2018). However, there exists a smoking-specific VSL study by Viscusi & Hersch (2007) who used the variation in smoking rates and fatality rates over the life-cycle to estimate the mortality costs (in 2000 US dollars) of smoking to be $1,538,631 for males and $563,299 for females, respectively. Taking the average of these values, we find the mortality cost of smoking to be $1,050,965 using the VSL estimates, which is equivalent to $1,628,635 in 2021 US dollars. We then discount the average present value mortality cost from age 58 to age 3: $1,628,635/ (1.03^{55}) = $320,462. So, the estimated mortality cost savings from a 5.8 percentage point reduction in smoking is expected to be $18,587.

Adding together the present values of reduced costs of medical expenditure, lost productivity, and premature mortality attributable to smoking, we estimate the present value of the effect of participation in the CPC program on the prevalence of smoking to be in the range of $7436 to $24,086.

Reduction in quality of life due to smoking. We also estimate the economic impact of smoking using the QALY method. In a recently published study, Xu *et al.* (2021) estimated that cigarette smokers aged 35–39 years lost an average of 7.4 QALYs as compared to non-smokers. Using the monetary value of $50,000 per QALY as in the section above, we estimate the savings at age 37 discounted to the present value at age 3 to be $7855.[6] Adding the additional healthcare costs attributable to smoking of $2735 as obtained above, the QALY approach yields a benefit of $10,591 from a reduced rate of smoking due to preschool participation.

As above for diabetes, we use the more conservative estimates of $7437 as the benefit from reduced smoking in our cost-benefit analysis. In Table 7, we report the benefits by category, and in Table 8, we report the total health benefits obtained using a variety of methods of estimation as described above.

7.3 Sensitivity analysis

There are various sources of uncertainty in our analysis, such as those associated with the coefficients of estimated health benefits from participating in the CPC program, discount rates used in our calculation, and the dollar estimate of benefits from the reduction in adverse health outcomes. We tested the robustness of economic benefits by using alternate

---

[6] (7.4 QALYs × $50,000 × 5.8%)/(1.03^34).

Table 7. *Distribution of benefits by category.*

Benefits	Diabetes	Smoking
Increased healthcare costs	5584	2735
Reduced productivity	1321	2764
Premature mortality		
Using foregone earnings method	554	1937
Using value of statistical life method	—	18,587
Reduction in QALYs	24,134	7855
Reduction in utility value	5076	—

Note: All benefit estimates are in 2021 US dollars.

discount rates and conducted a Monte Carlo simulation to estimate the likelihood of getting positive returns on investment considering the uncertainty in impact estimates and economic values.

7.3.1 Alternate discount rates

The discount rate reflects the worth of the dollar in the distant future as compared to the present. Since the choice of an appropriate discount rate can have an impact on the economic estimates, we used alternative discount rates of 1, 5, and 7 % as part of sensitivity analysis. Table 9 and Figure 3 compare the estimates of various cost and benefit components and the associated benefit-cost ratios and the net benefits for each of the discount rate values. Because the benefits are accrued much later in life, the net benefits of the program decrease with the increase in discount rates.

7.4 Monte Carlo analysis

We conducted a Monte Carlo analysis to account for uncertainty in the estimated effect of the program on outcomes. We ran 10,000 simulations based on the assumption of a normal distribution of the effect size and accounting for a 10 % margin in the dollar estimate of the health benefits. We found the benefit-cost ratio to be in the range of 0.30–2.72 with a mean of 1.36. We also find that in ~77 % of the simulations, the benefit-cost ratio was greater than one, indicating a positive economic return.

8 Discussion

In this study, we follow a cohort of students born in 1980 who attended kindergarten in high-poverty neighborhoods of Chicago, nearly two-thirds of who also attended a center-based preschool program. Making use of extensive information collected over more than three decades from the study participants, we evaluate the impact of the preschool program on the long-term health outcomes in adulthood.

We employ the inverse probability weighting approach to ensure comparability of the program and comparison groups and address missing data concerns and conduct a benefit-cost analysis. While there may be threats to internal validity due to non-observables

Table 8. *Total health benefits and benefit-cost ratio by estimation method.*

Benefits included	Estimate 1[a]	Estimate 2[b]	Estimate 3[c]	Estimate 4[d]
Healthcare costs	✓	✓	✓	✓
Lost productivity	✓	✓		
Premature mortality (using foregone earnings method)	✓			
Premature mortality (using VSL method for smoking)		✓		
QALYs lost			✓	
Lost utility value (for diabetes estimate)				✓
Total health benefits	$14,896	$31,546	$40,308	$21,250
Total costs	$11,000	$11,000	$11,000	$11,000
Benefit-cost ratio	1.35	2.87	3.66	1.93
Net-benefit	$3896	$20,546	$29,308	$10,250

[a]In this estimate, we add the healthcare costs, lost productivity costs, and premature mortality costs calculated through the foregone earning method for both diabetes and smoking.
[b]In this estimate, we add the premature mortality costs calculated through the foregone earning method for diabetes and VSL method for smoking to the healthcare costs and lost productivity costs.
[c]In this estimate, we add the healthcare costs to the economic cost of QALYs lost due to diabetes and smoking.
[d]In this estimate, we add the healthcare costs to the economic cost of QALYs lost due to smoking and lost utility value due to living with diabetes.

***Table 9.*** *Economic estimates based on the discount rate.*

Component	1 %	3 %	5 %	7 %
Program cost ($)	11,140	11,000	10,867	10,738
Benefits[a]				
Diabetes ($)	17,714	7459	3364	1613
Smoking ($)	16,279	7437	3751	2067
Total health benefits ($)	33,993	14,896	7115	3680
Benefit-cost ratio	3.05	1.35	0.65	0.34
Net benefit ($)	22,853	3896	−3752	−7058

[a]We use the most conservative estimates of diabetes and smoking benefits in this table.

predicting both participation and the outcome, we showed that the weights work well to ensure covariate balance. Our results indicate that participation in the CPC program is associated with a significantly lower prevalence of diabetes (effect size[7]: −0.17) and smoking (effect size: −0.31) in adulthood as well as lower body mass index. The effect size may be interpreted as small to medium, but they are nonetheless practically important considering that the effects are sustained over three decades (Kraft, 2020). In our study, we

---

[7] Probit effect size estimates were calculated for the dichotomous outcome variables.
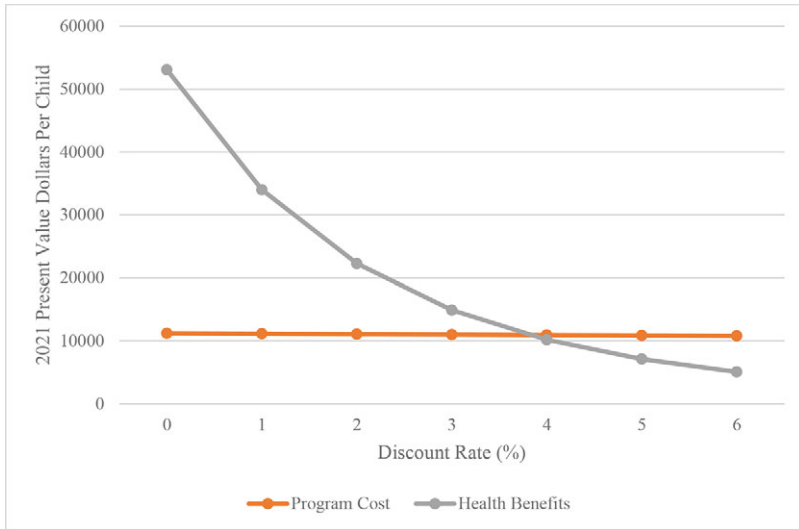
Figure 3. *Sensitivity analysis for estimated costs and health benefits for the CPC program.*

did not find any significant impacts of the program on depression, which could be attributed to the self-reported measure of diagnosis, as opposed to a multiple-item measure of symptoms. Indeed, a recent study by Mondi *et al.* (2020) found that the CPC program was associated with a 7.1 percentage point reduction in one or more depressive symptoms.

   Based on the calculations in the previous section, we obtain the total health benefits to be in the range of $14,896–40,308 depending on the preferred method of estimation of program benefits. Dividing by the cost of the program, which is estimated to be $11,000, we obtain a range of benefit-cost ratios from 1.35 to 3.66. A benefit-cost ratio of greater than one indicates that the health benefits of the program by themselves outweigh its costs. On subtracting the cost of the program from the total health benefits, we obtain the net benefit of the program to be in the range of $3896–29,308. While the health benefits of $1.35 per dollar invested are smaller than benefits from savings in crime reduction which were $4.99 per dollar invested, and benefits from increased earnings and tax revenues which were $3.39, they are higher than benefits from the reduction in spending on child welfare which was $0.86, and on special education, which was estimated to be $0.62 per dollar invested (Reynolds *et al.,* 2011*b*).

   We also conducted a Monte Carlo analysis to account for uncertainty in the estimated effect of the program on outcomes. This is important for two reasons; one, our estimates of the impact of the program on the health benefits could be imprecise, and second, the dollar value of benefits we use from the literature could vary based on the assumptions. From a simulation of 10,000 iterations, we found the Benefit-Cost Ratio to be in the range of 0.3–2.72 with the average ratio coming to 1.36 and a ratio of greater than one, in 77 % of the simulations.

   This study is strengthened by its use of correction for multiple hypothesis testing which reduces the false-discovery rates of multiple outcomes back to the level of an individual significance test. This procedure is rarely adopted in the literature while testing multiple

outcomes at once but is important to ensure that the statistical significance of each outcome is not achieved merely by chance. In addition to the range of methodologies employed for robustness check, another major strength of this study which adds to the validity of the results is its low attrition, with nearly three-fourth of the original participants still being followed-up after over 30 years, which is uncharacteristic of studies with a large sample greater than 1500.

On the lines of previous findings by Belfield *et al.* (2006), Muennig *et al.* (2009), and García *et al.* (2020) in their cost-benefit analyses of the Perry Preschool, as well as a previous cost-benefit analysis of the CPC intervention (Reynolds *et al.*, 2011a,b; Temple & Reynolds, 2015), the results suggest that the health impacts of early educational intervention are significant and may by themselves offset the costs of the intervention, even if no other benefits were observed.

Evidence that participation in school-based education intervention from preschool to third grade can have an impact on health outcomes in adulthood is of great importance to the current debate of how much resources to devote to publicly funded preschool and early elementary school programs. While most of the longer-term research evidence on the impacts of early education focus on outcomes such as earnings and involvement in criminal activity (e.g., Heckman *et al.,* 2010), this study adds to a much smaller set of recent studies on the potential for early investments in education to have long-lasting health effects. Certainly, some of the health benefits found in the CPC intervention are driven by the intervention's effect on educational attainment. While crime and violence reduction has been an important contributor to the large social benefits calculated for high-quality preschool investments in studies based on the Perry Preschool program and the CPC program, the new results reported in this article are likely to further amplify the rationale for government investments in early education by focusing more attention on the significant and economically important health savings that may accrue from these public investments.

9 Limitations and future directions

In this partial benefit-cost analysis, we focused solely on the health outcomes at age 37 for the preschool component of the CPC program. The results suggest that the health impacts of early educational intervention are significant and may by themselves offset the costs of the intervention, even if no other benefits were observed. However, a future study may look at incorporating benefits across a domain of outcomes such as gain in income and reduction in crime, in addition to health. This will help in calculating a comprehensive benefit-cost ratio of the program. A limitation of our study is that we rely on the self-reported data of health outcomes by the participants. However, the in-person health examination of CLS participants reported moderate to high correlations between self-reports and exams for BMI, hypertension, and diabetes (Reynolds, Eales *et al.*, 2021; Reynolds, Ou *et al.*, 2021). The Framingham Risk Score, which includes obesity and hypertension, in a broader index, provides complementary outcome data for further analysis. A future study could look at the health data collected through physical health exams which could enhance the range of outcomes studied and improve the validity of the outcomes. Some recent studies have also gone on to evaluate the impact of early childhood education programs such as the Perry preschool program on mortality (Heckman & Karapakula, 2019). Further research would be needed to evaluate the impact of the CPC program on the premature mortality of participants.

In a prior benefit-cost analysis of this sample, a small induced cost (or negative benefit) was reported in terms of the higher required costs to taxpayers if preschool participation induces greater enrollment in public universities (Reynolds *et al.*, 2011*b*). However, this current health study does not report additional costs or negative benefits arising from better health and longevity. Higher expenditures on social security retirement benefits might be an example of such an induced cost.

While the estimation of the effects of program participation on the various health outcomes suggested significant effects on diabetes, smoking, and BMI, our corrections for multiple inferences led to these results holding at a seemingly high false discovery rate of 0.2. McDonald (2009) and Lee & Lee (2018) have recommended using an FDR value of 0.2 when the cost of false positives (missing potentially important findings) is high. However, if we were to use a more conservative FDR value of 0.1 or lower, then these estimates of program impact would no longer be statistically significant at conventional levels. Finally, we only examined depressive symptoms as a measure of mental health outcomes. Mondi *et al.* (2020) examined a broader measure of midlife psychological well-being in which CPC participants fared better. Other indicators such as life satisfaction warrant further investigation and potential inclusion in economic returns. Future studies in this field also can increase the focus on mental health by studying other determinants of mental health. This will help us get a better understanding of the impact on overall health including physical and mental health.

Competing Interests. The authors declare none.

References

American Diabetes Association. 2018. "Economic Costs of Diabetes in the U.S. in 2017." *Diabetes Care*, *41*(5): 917–928. https://doi.org/10.2337/dci18-0007.

Austin, P. C., M. M. Mamdani, T. A. Stukel, G. M. Anderson, and J. V. Tu. 2005. "The Use of the Propensity Score for Estimating Treatment Effects: Administrative Versus Clinical Data." *Statistics in Medicine*, *24*(10): 1563–1578.

Austin, P. C. and E. A. Stuart. 2015. "Moving Towards Best Practice when using Inverse Probability of Treatment Weighting (IPTW) Using the Propensity Score to Estimate Causal Treatment Effects in Observational Studies." *Statistics in Medicine*, *34*(28): 3661–3679. https://doi.org/10.1002/sim.6607.

Bailey, M. J., B. D. Timpe, and S. Sun. 2020. *Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency*. Cambridge, MA: National Bureau of Economic Research.

Barnett, W. S. 1996. *Lives in the Balance: Age-27 Benefit-Cost Analysis of the High/Scope Perry Preschool Program. Monographs of the High/Scope Educational Research Foundation, Number Eleven.* ERIC.

Barnett, W. S., and L. N. Masse. 2007. "Comparative Benefit-Cost Analysis of the Abecedarian Program and its Policy Implications." *Economics of Education Review*, *26*(1): 113–125. https://doi.org/10.1016/j.econedurev.2005.10.007.

Barr, A. C., and C. Gibbs. 2019. *Breaking the Cycle? Intergenerational Effects of an Anti-Poverty Program in Early Childhood* (Working Paper No. 19-141). Available at http://www.edworkingpapers.com/ai19-141 (accessed June 08, 2021).

Bauer, L. 2019. *Does Head Start work? The Debate Over the Head Start Impact Study, Explained*. Washington, DC: Brookings. https://doi.org/10.1007/S11292-013-9190-3.PDF.

Bauer, L., and D. W. Schanzenbach 2016. *The Long-Term Impact of the Head Start Program.* Available at https://www.hamiltonproject.org/assets/files/long_term_impact_of_head_start_program.pdf (accessed March 09, 2020).

Beaudet, A., J. Clegg, P.-O. Thuresson, A. Lloyd, and P. McEwan. 2014. "Review of Utility Values for Economic Modeling in Type 2 Diabetes." *Value in Health*, *17*(4): 462–470.

84   *Nishank Varshney et al.*

Belfield, C. R., M. Nores, W. S. Barnett, and L. J. Schweinhart. 2006. "The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup." *The Journal of Human Resources*, *41*(1): 162–190. Available at http://www.jstor.org/stable/40057261 (accessed December 08, 2021).

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1): 289–300.

Camilli, G., S. Vargas, S. Ryan, and W. S. Barnett. 2010. "Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development." *Teachers College Record*, *112*(3): 579–620.

Campbell, F. A., G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, E. P. Pungello, and Y. Pan. 2014. "Early Childhood Investments Substantially Boost Adult Health." *Science*, *343*(6178): 1478–1485.

Campbell, F. A., E. P. Pungello, M. Burchinal, K. Kainz, Y. Pan, B. H. Wasik, O. A. Barbarin, J. J. Sparling, and C. T. Ramey. 2012. "Adult Outcomes as a Function of an Early Childhood Educational Program: An Abecedarian Project Follow-Up." *Developmental Psychology*, *48*(4): 1033–1043. https://doi.org/10.1037/a0026644.

Campbell, F. A., C. T. Ramey, E. P. Pungello, J. J. Sparling, and S. Miller-Johnson. 2002. "Early Childhood Education: Young Adult Outcomes from the Abecedarian Project." *Applied Developmental Science*, *6*(1): 42–57.

Cannon, J. S., M. R. Kilburn, L. A. Karoly, T. Mattox, A. N. Muchow, and M. Buenaventura. 2017. "Investing Early." In *Taking Stock of Outcomes and Economic Returns from Early Childhood Programs*. Santa Mónica, CA: RAND Corporation.

Cannon, J. S., M. R. Kilburn, L. A. Karoly, T. Mattox, A. N. Muchow, and M. Buenaventura. 2018. "Investing Early: Taking Stock of Outcomes and Economic Returns from Early Childhood Programs." *Rand Health Quarterly*, *7*(4): 6.

Carneiro, P., and R. 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. https://doi.org/10.1257/pol.6.4.135.

Centers for Disease Control and Prevention. 2005. *Cigarette Smoking Among Adults - United States, 2004.* Available at https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5444a2.htm (accessed February 08, 2020).

Centers for Disease Control and Prevention. 2008. *Smoking-Attributable Mortality, Years of Potential Life Lost, and Productivity Losses - United States, 2000–2004.* Available at https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5745a3.htm (accessed February 08, 2020).

Centers for Disease Control and Prevention. 2017. *Table 15. Life Expectancy at Birth, at Age 65, and at Age 75, by Sex, Race, and Hispanic Origin: United States, Selected Years 1900–2016.* Available at https://www.cdc.gov/nchs/hus/contents2017.htm#015 (accessed November 01, 2021).

Clarke, P., A. Gray, and R. Holman. 2002. "Estimating Utility Values for Health States of Type 2 Diabetic Patients using the EQ-5D (UKPDS 62)." *Medical Decision Making*, *22*(4): 340–349.

Council of Economic Advisers. 2016. *The Disconnect Between Resources and Needs When Investing in Children.*

Crowley, D. M., K. A. Dodge, W. S. Barnett, P. Corso, S. Duffy, P. Graham, M. Greenberg, R. Haskins, L. Hill, and D. E. Jones. 2018. "Standards of Evidence for Conducting and Reporting Economic Evaluations in Prevention Science." *Prevention Science*, *19*(3): 366–390.

Currie, J., and Thomas, D. 1995. "Does Head Start Make a Difference?" In *The American Economic Review* (Vol. *85*): 341–364. Nashville, TN: American Economic Association. https://doi.org/10.2307/2118178.

Deming, D. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics*, *1*(3): 111–134. https://doi.org/10.1257/app.1.3.111.

Dietrichson, J., I. L. Kristiansen, and B. C. V. Nielsen. 2018. *Universal Preschool Programs and Long-Term Child Outcomes: A Systematic Review*. Working Paper.

Eales, L., A. J. Reynolds, and S.-R. Ou. 2020. "Childhood Predictors of Adult Obesity in the Chicago Longitudinal Study." *Preventive Medicine*, *132*: 105993.

Englund, M. M., B. A. B. White, A. J. Reynolds, L. J. Schweinhart, and F. A. Campbell. 2014. "Health Outcomes of the Abecedarian, Child-Parent Center, and HighScope Perry Preschool Programs." In Reynolds, A. I., A. T. Rolnick, and J. A. Temple (Eds.) *Health and Education in Early Childhood: Predictors, Interventions, and Policies* (Vol. *2014*): 257–292. Cambridge: Cambridge University Press.

Feller, A., T. Grindal, L. Miratrix, and L. C. Page. 2016. "Compared to What? Variation in the Impacts of Early Childhood Education by Alternative Care Type." *The Annals of Applied Statistics*, *10*(3): 1245–1285.

Friedman-Krauss, A. H., W. S. Barnett, K. A. Garver, K. S. Hodges, G. G. Weisenfeld, and N. DiCrecchio. 2019. *The State of Preschool 2018: State Preschool Yearbook*. New Brunswick, NJ: National Institute for Early Education Research.

Garces, E., D. Thomas, and J. Currie. 2002. "Longer-Term Effects of Head Start." In *The American Economic Review* (Vol. *92*): 999–1012. Nashville, TN: American Economic Association. https://doi.org/10.2307/3083291.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados. 2020. "Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program." *Journal of Political Economy*, *128*(7): 2502–2541.

García, J. L., J. J. Heckman, and A. L. Ziff. 2018. "Gender Differences in the Benefits of an Influential Early Childhood Program." *European Economic Review*, *109*: 9–22.

Grosse, S. D. 2008. "Assessing Cost-Effectiveness in Healthcare: History of the \$50,000 Per QALY Threshold." *Expert Review of Pharmacoeconomics & Outcomes Research*, *8*(2): 165–178.

Heckman, J. J. 2008. "Schools, Skills, and Synapses." *National Bureau of Economic Research*, *46*(3): 289–324. https://doi.org/10.3386/w14064.

Heckman, J. J. and G. Karapakula. 2019. *The Perry Preschoolers at Late Midlife: A Study in Design-Specific Inference*. Cambridge, MA: National Bureau of Economic Research.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics*, *94*(1–2): 114–128.

Hirano, K., G. W. Imbens, and G. Ridder. 2003. "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score." *Econometrica*, *71*(4): 1161–1189.

Huang, E. S., S. E. S. Brown, B. G. Ewigman, E. C. Foley, and D. O. Meltzer. 2007. "Patient Perceptions of Quality of Life with Diabetes-Related Complications and Treatments." *Diabetes Care*, *30*(10): 2478–2483.

Jha, P., C. Ramasundarahettige, V. Landsman, B. Rostron, M. Thun, R. N. Anderson, T. McAfee, and R. Peto. 2013. "21st-Century Hazards of Smoking and Benefits of Cessation in the United States." *New England Journal of Medicine*, *368*(4): 341–350. https://doi.org/10.1056/NEJMsa1211128.

Karoly, L. A. 2012. Toward Standardization of Benefit-Cost Analysis of Early Childhood Interventions. *Journal of Benefit-Cost Analysis*, *3*(1), 1–45. https://doi.org/10.1515/2152-2812.1085.

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

Knight, D. S., S. Landry, T. A. Zucker, E. C. Merz, C. L. Guttentag, and H. B. Taylor. 2019. "Cost-Effectiveness of Early Childhood Interventions to Enhance Preschool: Evidence from a Randomized Experiment in Head Start Centers Enrolling Historically Underserved Populations." *Journal of Policy Analysis and Management*, *38*(4): 891–917.

Kraft, M. A. 2020. "Interpreting Effect Sizes of Education Interventions." *Educational Researcher*, *49*(4): 241–253.

Lee, S. and D. K. Lee. 2018. "What is the Proper Way to Apply the Multiple Comparison Test?" *Korean Journal of Anesthesiology*, *71*(5): 353.

Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2017). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. SAGE publications.

Lipsey, M. W., D. C. Farran, and K. Durkin. 2018. "Effects of the Tennessee Prekindergarten Program on Children's Achievement And Behavior Through Third Grade." *Early Childhood Research Quarterly*, *45*: 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005.

McCoy, D. C., H. Yoshikawa, K. M. Ziol-Guest, G. J. Duncan, H. S. Schindler, K. Magnuson, R. Yang, A. Koepp, and J. P. Shonkoff. 2017. "Impacts of Early Childhood Education on Medium-and Long-Term Educational Outcomes." *Educational Researcher*, *46*(8): 474–487.

McDonald, J. H. 2009. *Handbook of Biological Statistics* (Vol. *2*). Baltimore, MD: Sparky House.

Mondi, C. F., A. J. Reynolds, and B. A. Richardson. 2020. "Early Childhood Education and Adult Depression: An Attrition Reanalysis with Inverse Propensity Score Weighting." *Evaluation Review*, *44*(5–6): 379–409.

Muennig, P., D. Robertson, G. Johnson, F. Campbell, E. P. Pungello, and M. Neidell. 2011. "The Effect of an Early Education Program on Adult Health: The Carolina Abecedarian Project Randomized Controlled Trial." *American Journal of Public Health*, *101*(3): 512–516.

Muennig, P., L. J. Schweinhart, J. Montie, and M. Neidell. 2009. "Effects of a Prekindergarten Educational Intervention on Adult Health: 37-Year Follow-Up Results of a Randomized Controlled Trial." *American Journal of Public Health*, *99*(8): 1431–1437.

Narayan, K. M. V., J. P. Boyle, T. J. Thompson, S. W. Sorensen, and D. F. Williamson. 2003. "Lifetime Risk for Diabetes Mellitus in the United States." *Journal of the American Medical Association*, *290*(14): 1884–1890.

Neumann, A., O. Schoffer, F. Norström, M. Norberg, S. J. Klug, and L. Lindholm. 2014. "Health-Related Quality of Life for Pre-Diabetic States and Type 2 Diabetes Mellitus: A Cross-Sectional Study in Västerbotten Sweden." *Health and Quality of Life Outcomes*, *12*(1): 1–10.

Nores, M., C. R. Belfield, W. S. Barnett, and L. Schweinhart. 2005. "Updating the Economic Impacts of the High/Scope Perry Preschool Program." *Educational Evaluation and Policy Analysis*, *27*(3): 245–261.

Ou, S.-R., C. F. Mondi, S. Yoo, K. Park, B. Warren, and A. J. Reynolds. 2020. "Thirty Years Later: Locating and Interviewing Participants of the Chicago Longitudinal Study." *Early Childhood Research Quarterly*, *51*: 1–13.

Pearman, F. A., Springer, M., M. W. Lipsey, M. Lachowicz, W. Swain, and D. C. Farran. 2020. "Teachers, Schools, and Pre-K Effect Persistence: An Examination of the Sustaining Environment Hypothesis." *Journal of Research on Educational Effectiveness*, *13*(4):547–573.

Peterson, C., C. Florence, and J. Klevens. 2018. "The Economic Burden of Child Maltreatment in the United States, 2015." *Child Abuse & Neglect*, *86*: 178–183.

Puma, M., S. Bell, R. Cook, C. Heid, P. Broene, F. Jenkins, A. Mashburn, and J. Downer. 2012. *Third Grade Follow-Up to the Head Start Impact Study: Final Report. OPRE Report 2012–45. Administration for Children & Families.*

Puma, M., S. Bell, R. Cook, C. Heid, and M. Lopez. 2005. *Head Start Impact Study: First Year Findings. Administration for Children & Families.*

Puma, M., S. Bell, R. Cook, C. Heid, G. Shapiro, P. Broene, F. Jenkins, P. Fletcher, L. Quinn, and J. Friedman. 2010. *Head Start Impact Study. Final Report. Administration for Children & Families.*

Ramon, I., S. K. Chattopadhyay, W. S. Barnett, and R. A. Hahn. 2018. "Early Childhood Education to Promote Health Equity: A Community Guide Economic Review." *Journal of Public Health Management and Practice*, *24*(1): e8–e15.

Reynolds, A. J. 2000. *Success in Early Intervention: The Chicago Child Parent Centers*. Lincoln, NE: University of Nebraska Press.

Reynolds, A. J. and S.-R. Ou. 2016. "Generative Mechanisms in Early Childhood Interventions: A Confirmatory Research Framework for Prevention." *Prevention Science*, *17*(7): 794–805.

Reynolds, A. J., S.-R. Ou, and J. A. Temple. 2018. "A Multicomponent, Preschool to Third Grade Preventive Intervention and Educational Attainment at 35 Years of Age." *JAMA Pediatrics*, *172*(3): 247–256.

Reynolds, A. J., S.-R. Ou, and J. W. Topitzes. 2004. "Paths of Effects of Early Childhood Intervention on Educational Attainment and Delinquency: A Confirmatory Analysis of the Chicago Child-Parent Centers." *Child Development*, *75*(5): 1299–1328.

Reynolds, A. J., S.-R. Ou, C. F. Mondi, and A. Giovanelli. 2019. "Reducing poverty and inequality through preschool-to-third-grade prevention services." *American Psychologist*, *74*(6): 653–672.

Reynolds, A. J. and J. A. Temple. 2019. *Sustaining Early Childhood Learning Gains: Program, School, and Family Influences*. Cambridge: Cambridge University Press.

Reynolds, A. J., J. A. Temple, and S.-R. Ou. 2003. "School-Based Early Intervention and Child Well-Being in the Chicago Longitudinal Study." *Child Welfare*, *82*(5): 633–656.

Reynolds, A. J., J. A. Temple, S.-R. Ou, I. A. Arteaga, and B. A. B. White. 2011a. "School-Based Early Childhood Education and Age-28 Well-Being: Effects by Timing, Dosage, and Subgroups." *Science*, *333*(6040): 360–364.

Reynolds, A. J., J. A. Temple, D. L. Robertson, and E. A. Mann. 2001. "Long-Term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest: A 15-Year Follow-Up of Low-Income Children in Public Schools." *JAMA*, *285*(18): 2339–2346.

Reynolds, A. J., J. A. Temple, D. L. Robertson, and E. A. Mann. 2002. "Age 21 Cost-Benefit Analysis of the Title I Chicago Child-Parent Centers." *Educational Evaluation and Policy Analysis*, *24*(4): 267–303.

Reynolds, A. J., J. A. Temple, B. A. B. White, S.-R. Ou, and D. L. Robertson. 2011b. "Age 26 Cost–Benefit Analysis of the Child-Parent Center Early Education Program." *Child Development*, *82*(1): 379–404.

Reynolds, A. J., Eales, L., Ou, S. R., Mondi, C. F., & Giovanelli, A. (2021). A comprehensive, multisystemic early childhood program and obesity at age 37 years. *JAMA pediatrics*, *175*(6), 637–640.

Reynolds, A. J., Ou, S. R., Eales, L., Mondi, C. F., & Giovanelli, A. (2021). Assessment of a Comprehensive Early Childhood Education Program and Cardiovascular Disease Risk in Midlife. *JAMA Network Open*, *4*(8), e2120752–e2120752.

Rhodes, E. T., L. A. Prosser, T. J. Hoerger, T. Lieu, D. S. Ludwig, and L. M. Laffel. 2012. "Estimated Morbidity and Mortality in Adolescents and Young Adults Diagnosed with Type 2 Diabetes Mellitus." *Diabetic Medicine*, *29* (4): 453–463.

Schochet, P. Z. 2008. *Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations. NCEE 2008-4018.* National Center for Education Evaluation and Regional Assistance.

Schulte, B., and A. Durana. 2016. *The New America Care Report. New America Policy Paper.*

Schweinhart, L. J. 1993. *Significant Benefits: The High/Scope Perry Preschool Study through Age 27. Monographs of the High/Scope Educational Research Foundation, No. Ten.* ERIC.

Schweinhart, L. J. 2013. "Long-Term Follow-Up of a Preschool Experiment." *Journal of Experimental Criminology*, 9(4): 389–409. https://doi.org/10.1007/s11292-013-9190-3.

Schweinhart, L. J., and D. P. Weikart. 1997. "The High/Scope Preschool Curriculum Comparison Study Through Age 23." *Early Childhood Research Quarterly*, 12(2): 117–144.

Seaman, S. R., and I. R. White. 2013. "Review of Inverse Probability Weighting for Dealing with Missing Data." *Statistical Methods in Medical Research*, 22(3): 278–295. https://doi.org/10.1177/0962280210395740.

Shiroiwa, T., T. Fukuda, S. Ikeda, and K. Shimozuma. 2013. "QALY and Productivity Loss: Empirical Evidence for "Double Counting"." *Value in Health*, 16(4): 581–587.

Shonkoff, J. P. 2010. "Building a New Biodevelopmental Framework to Guide the Future of Early Childhood Policy." *Child Development*, 81(1): 357–367.

Temple, J. A., and A. J. Reynolds. 2015. "Using Benefit-Cost Analysis to Scale up Early Childhood Programs through Pay-for-Success Financing." *Journal of Benefit-Cost Analysis*, 6(3): 628–653.

Topitzes, J. W., O. Godes, J. P. Mersky, S. Ceglarek, and A. J. Reynolds. 2009. "Educational Success and Adult Health: Findings from the Chicago Longitudinal Study." *Prevention Science*, 10(2): 175–195.

Trogdon, J. G., and T. Hylands. 2008. "Nationally Representative Medical Costs of Diabetes by Time since Diagnosis." *Diabetes Care*, 31(12): 2307–2311.

U.S. Department of Health and Human Services. 2014. *The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General*. Atlanta, GA: US Department of Health and Human Services.

US Census Bureau. 2017. *More Than Half of School Expenditures Spent on Classroom Instruction.* Available at https://www.census.gov/newsroom/press-releases/2017/cb17-97-public-education-finance.html (accessed March 10, 2020).

van Huizen, T., L. Dumhs, and J. Plantenga. 2019. "The Costs and Benefits of Investing in Universal Preschool: Evidence from a Spanish Reform." *Child Development*, 90(3): e386–e406.

Viscusi, W. K., and Hersch, J. 2007. *The Mortality Cost to Smokers.* Available at http://www.nber.org/papers/w13599 (accessed March 09, 2019).

Xu, X., L. Fiacco, B. Rostron, G. Homsi, E. Salazar, B. Levine, C. Ren, and J. Nonnemaker. 2021. "Assessing Quality-Adjusted Years of Life Lost Associated with Exclusive Cigarette Smoking and Smokeless Tobacco Use." *Preventive Medicine*, 150: 106707.

Zhang, P., M. B. Brown, D. Bilik, R. T. Ackermann, R. Li, and W. H. Herman. 2012. "Health Utility Scores for People with Type 2 Diabetes in US Managed Care Health Plans: Results from Translating Research Into Action for Diabetes (TRIAD)." *Diabetes Care*, 35(11): 2250–2256.

Appendix

Table 10. *Variables included as covariates in the regression and inverse probability weighting (IPW) models for program selection and attrition.*

Variable	IPW selection model	IPW attrition model	Regression model
Mother did not complete HS, child age 0–3	Yes	Yes	Yes
Child eligible for subsidized meals, child age 0–3	Yes	Yes	Yes
Mother under age 18 at childbirth	Yes	Yes	Yes
Four or more children in the family, child age 0–3	Yes	Yes	Yes
Participate in AFDC program, child age 0–3	Yes	Yes	Yes
Mother not employed, child age 0–3	Yes	Yes	Yes
Single parent family status, child age 0–3	Yes	Yes	Yes
Reside in high poverty neighborhood	Yes	Yes	Yes
Indicator for missing risk factors, child age 0–3	Yes	Yes	Yes
Family conflict, child age 0–5	Yes	Yes	Yes
Family financial problems, child age 0–5	Yes	Yes	Yes
Substance abuse parent, child age 0–5	Yes	Yes	Yes
Female child	Yes	Yes	Yes
African American child	Yes	Yes	Yes
CPC preschool program participation	No	Yes	Yes
CPC school-age program participation	No	Yes	Yes
Standardized word test, child age 5	No	Yes	No
Proxy of residential mobility	No	Yes	No
Have social security number	No	Yes	No
Census tract neighborhood mobility <1 year	No	Yes	No
Census tract neighborhood mobility 1–5 years	No	Yes	No
Census tract neighborhood mobility 5–10 years	No	Yes	No
Census tract neighborhood mobility 10–20 years	No	Yes	No
Census tract self-employed rate	No	Yes	No
Census tract African American female householder	No	Yes	No

Note: Data on these variables were collected from school administrative records and parent surveys.

Table 11. *IPW adjusted regression outcomes.*

Variables	(1) Outcome 1: Smoking	(2) Outcome 2: Drug Use	(3) Outcome 3: BMI	(4) Outcome 4: Obesity	(5) Outcome 5: Diabetes	(6) Outcome 6: Hypertension	(7) Outcome 7: Depression
CPC	−0.058*	−0.025	−0.960*	−0.047	−0.041**	−0.000	0.011
	(0.031)	(0.018)	(0.528)	(0.035)	(0.018)	(0.027)	(0.025)
CPC School-age program participation	0.018	0.005	0.110	0.002	0.011	−0.009	0.028
	(0.031)	(0.018)	(0.529)	(0.035)	(0.017)	(0.027)	(0.025)
Female child	−0.086***	−0.086***	2.971***	0.150***	0.026	0.025	−0.055**
	(0.028)	(0.016)	(0.482)	(0.033)	(0.017)	(0.026)	(0.022)
African-American child	0.058	−0.022	0.177	−0.006	0.051**	0.012	0.049
	(0.065)	(0.045)	(0.811)	(0.074)	(0.022)	(0.062)	(0.049)
Mother did not complete high school	0.009	0.023	−0.724	−0.025	0.033*	0.045*	0.046*
	(0.031)	(0.018)	(0.514)	(0.036)	(0.017)	(0.027)	(0.025)
Child eligible for subsidized meals	0.022	0.003	0.275	0.044	0.019	−0.045	0.009
	(0.038)	(0.022)	(0.675)	(0.047)	(0.019)	(0.036)	(0.027)
Mother under age 18 at childbirth	0.048	−0.035	0.630	0.006	−0.005	0.007	−0.049
	(0.046)	(0.024)	(0.674)	(0.050)	(0.025)	(0.042)	(0.034)
Four or more children in the family	0.047	0.002	1.038	0.026	−0.013	−0.001	0.045
	(0.040)	(0.023)	(0.675)	(0.045)	(0.022)	(0.033)	(0.031)
Participate in AFDC program	0.029	0.003	−0.611	−0.041	−0.026	0.035	0.047
	(0.038)	(0.018)	(0.660)	(0.048)	(0.027)	(0.035)	(0.029)
Mother not employed	0.037	−0.005	0.719	0.048	0.020	−0.022	0.014
	(0.036)	(0.018)	(0.653)	(0.047)	(0.026)	(0.033)	(0.029)
Single parent family status	−0.031	−0.009	−0.187	−0.019	−0.023	0.052*	0.019
	(0.036)	(0.022)	(0.582)	(0.040)	(0.024)	(0.030)	(0.026)

Table 11. *Continued*

Variables	(1) Outcome 1: Smoking	(2) Outcome 2: Drug Use	(3) Outcome 3: BMI	(4) Outcome 4: Obesity	(5) Outcome 5: Diabetes	(6) Outcome 6: Hypertension	(7) Outcome 7: Depression
Indicator for missing	0.080*	0.037	−0.791	−0.091**	−0.012	0.005	0.007
risk factors	(0.042)	(0.027)	(0.636)	(0.045)	(0.021)	(0.035)	(0.032)
Reside in high poverty	0.033	−0.064***	0.294	−0.035	−0.006	0.028	−0.033
neighborhood	(0.031)	(0.022)	(0.498)	(0.037)	(0.019)	(0.028)	(0.026)
Family conflict	0.033	0.004	−2.152**	−0.141*	0.020	0.037	0.114*
	(0.072)	(0.045)	(0.889)	(0.074)	(0.033)	(0.059)	(0.063)
Family financial	−0.050	0.082*	1.616*	0.047	−0.010	−0.021	0.087
problems	(0.056)	(0.049)	(0.891)	(0.068)	(0.022)	(0.049)	(0.055)
Substance abuse parent	0.044	0.002	0.252	−0.049	−0.010	−0.008	−0.120**
	(0.074)	(0.049)	(1.173)	(0.082)	(0.034)	(0.070)	(0.047)
Constant	0.142**	0.173***	29.240***	0.441***	0.011	0.095	0.018
	(0.069)	(0.058)	(1.090)	(0.092)	(0.030)	(0.073)	(0.054)
Observations	1,100	1,097	1,065	1,065	1,097	1,096	1,098
R-squared	0.031	0.067	0.063	0.039	0.020	0.013	0.045

Statistical Significance Levels *** p<0.01, ** p<0.05, * p<0.1