NBER WORKING PAPER SERIES

DETECTING DRIVERS OF BEHAVIOR AT AN EARLY AGE: EVIDENCE FROM A LONGITUDINAL FIELD EXPERIMENT

Marco Castillo
John A. List
Ragan Petrie
Anya Samek

Working Paper 28288
http://www.nber.org/papers/w28288

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020

We thank Edie Dobrez, Kristin Troutman, Amanda Chuan, Rui Chen, Andre Gray and Shreemayi Samujjwala for excellent research assistance. This research was funded by the Kenneth and Anne Griffin Foundation and by NIH grant 1R01DK114238. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Detecting Drivers of Behavior at an Early Age: Evidence from a Longitudinal Field Experiment
Marco Castillo, John A. List, Ragan Petrie, and Anya Samek
NBER Working Paper No. 28288
December 2020
JEL No. C91,C93,D12,D81,I21,I26

ABSTRACT

We use field experiments with nearly 900 children to investigate how skills developed at ages 3-5 drive later-life outcomes. We find that skills map onto three distinct factors - cognitive skills, executive functions, and economic preferences. Returning to the children up to 7 years later, we find that executive functions, but not cognitive skills, predict the likelihood of receiving disciplinary referrals. Economic preferences have an independent effect: children who displayed impatience at ages 3-5 were more likely to receive disciplinary referrals. Random assignment to a parenting program reduced disciplinary referrals. This effect was not mediated by skills or preferences.

Marco Castillo
Texas A&M University
marco.castillo@tamu.edu

John A. List
Department of Economics
University of Chicago
1126 East 59th
Chicago, IL 60637
and NBER
jlist@uchicago.edu

Ragan Petrie
Department of Economics
Texas A&M University
4228 TAMU
College Station, TX 77843
and The Melbourne Institute
rpetrie@tamu.edu

Anya Samek
University of California, San Diego
Rady School of Management
Wells Fargo Hall
9500 Gilman Drive #0553
La Jolla, CA 92093-0553
and NBER
anyasamek@gmail.com

1 Introduction

In a landmark study in the 1970s, hundreds of 4-year-old children decided whether to eat a marshmallow now, or wait until later and eat two marshmallows. The research, now popularly dubbed the "marshmallow test," is a psychological experiment that measured whether the child could resist the temptation of eating the marshmallow immediately (Mischel et al., 1972). Reports suggested that children who delayed gratification in this task at the age of four had improved cognition, educational outcomes, and health in adolescence and adulthood (Mischel et al., 1989; Shoda et al., 1990; Schlam et al., 2013).[1]

The marshmallow test captured the imagination of the public and generated hundreds of academic citations in part because of humankind's desire to understand the driving forces that shape our behavior. In addition to delay of gratification, researchers point to the importance of executive functioning skills - e.g., working memory, inhibitory control, and attention (Zelazo et al., 2016). Educators, on the other hand, have historically paid attention to early life cognitive or academic kills - e.g., reading, writing, and math. Finally, a small but growing literature points to an important role of economic preferences - e.g., time, risk, and social preferences (List et al., 2018). Indeed, models of human capital formation suggest that early life skills are key predictors of later life outcomes, and recognize potential synergies among skills in the education production function (Carneiro and Heckman, 2003).

An understanding of the interrelationships between the above-mentioned skills and their contributions to later-life outcomes is a necessary ingredient for the optimal design of early education policy. There is now a consensus that early skills matter (Heckman,

[1]The interpretation that the marshmallow test only captured delay of gratification has come under some scrutiny. One study showed that varying environment reliability changed child wait time, suggesting that at least part of the ability to wait is driven by trust in the environment (Kidd et al., 2013). Another study showed that the delay of gratification data collected in early childhood did not predict mid-life outcomes and survey data collected in adolescence was more predictive (Benjamin et al., 2020).

2000; Carneiro and Heckman, 2003), but there is limited work exploring the interrelationships of different skills. One exception is Cunha et al. (2010), who use data from the National Longitudinal Survey of Youth to estimate the technology of the formation of cognitive ability and non-cognitive ability such as temperament and behavior.[2] Given the growing interest in investing in early childhood programs, such evidence is becoming increasingly important as it can help answer questions about how skills shape outcomes (Currie, 2001; Almond et al., 2018).

In this study, we conduct field experiments with a large group of children aged 3-5 years old to generate data on their cognitive skills (receptive vocabulary, letter recognition, writing and math), executive functions (working memory, inhibitory control, emotion regulation and attention), economic preferences (time, risk and social preferences) and delay of gratification.[3] The data are paired with survey data on demographic and socio-economic (SES) background to explore early-life correlations. This analysis allows us to evaluate how the critical skills of executive functioning and economic preferences correlate with cognitive skills, data that are not evaluated in Cunha et al. (2010).

A key novelty of our work is that we return to the same children when they are ages 5-12 and collect administrative data on disciplinary referrals at school, which are considered an important predictor of adult outcomes (e.g. Segal, 2013). Finally, we take advantage of a field experiment in which a sub-set of our participants were randomized into an early childhood program—which we conducted and designed—called the Chicago Heights Early Childhood Center (CHECC). The data complementarities permit us to evaluate the causal impact of early childhood programs on skills and disciplinary referrals while simultaneously exploring whether early life skills mediate the

[2]More recently, Heckman et al. (2020) evaluates the impact of home visiting programs in China on an array of skills, including language, executive functions and social-emotional skills.

[3]The literature most often refers to executive functions as 'higher order cognitive skills.' Such skills are rarely targeted by standard academic achievement tests that measure cognitive ability; hence, in this paper we consider them separately.

influences of CHECC.

We report four main results. Our first finding is that cognitive skills, executive functions, and economic preferences measured at an early age map to distinct 'factors.' Importantly, economic preferences are not strongly associated with cognitive skills or executive functions at this early age and map to their own factors.[4] In related work with children, Bettinger and Slonim (2007) find no association between time preferences and math skills, whereas Andreoni et al. (2019a) find an association between risk preferences and cognitive abilities. Adolescents and adults seem to display stronger relationships between skills and preferences (Burks et al., 2009; Benjamin et al., 2013). Taken together, these findings open up the distinct possibility that economic preferences may be less correlated with skills at early ages, but preferences drive decisions about human capital investments, such that preferences and skills assessed later in life are correlated.

Our second finding is that some - but not all - early life skills have substantial predictive power on later life outcomes. A statistically significant relationship exists for executive functions: an increase of one standard deviation in executive functions is associated with an 1.9 percentage point decrease in the likelihood of disciplinary referrals several years later. The baseline likelihood of disciplinary referral is 14.7%. On the other hand, cognitive skills measured at ages 3-5 do not predict disciplinary referrals in our data. The result that executive functions - which encompass attention and emotion regulation - drive future disciplinary referrals is consonant with Heckman (2000), who advances an argument pertaining to the importance of 'non-cognitive' or 'soft' skills that are not targeted by typical academic achievement tests.

Our third finding relates to the relationship between economic preferences and disciplinary referrals. We might have expected young children who are more patient,

[4]Interestingly, delay of gratification as measured by the marshmallow task does not map to any of the main factors that we identify and is not strongly associated with time preferences as measured by our experiment. The major difference between the two paradigms is that in the latter, we give children a deliberate choice between now and later, whereas in the former children need to exercise inhibitory control for a length of time to avoid eating the marshmallow.

less risk-seeking, and more pro-social to have fewer disciplinary referrals than their counterparts.[5] This expectation is borne out for time preferences: we find that a one standard deviation increase in patience as measured by our time preference task is associated with an 2.3 percentage point decrease in the likelihood of disciplinary referrals. This relationship holds even when controlling for a wealth of other factors, including academic skills and executive functions. Alternatively, risk preference, social preferences, or delay of gratification as measured by the marshmallow task do not have similar predictive power.

Our fourth finding is that random assignment to our early childhood program, CHECC, and in particular into CHECC's parenting program - which encouraged parents to teach to their children at home - reduced the likelihood of disciplinary referrals later in life. However, while these programs have some impact on the skills that we measure, we do not find evidence that changes in these skills mediate the impact of our CHECC treatments. We conclude that other unmeasured skills or mechanisms potentially drive the effects. In this spirit, our analysis joins the work of Heckman et al. (2013), who studied the mechanisms behind the long-term effects of the Perry Preschool project.

Overall, we view our work as speaking to several distinct literatures. We contribute to the literature on human capital development by providing data on the roles of cognitive skills, executive functions, and economic preferences. This literature has largely overlooked the role of economic preferences of children. Our finding that economic preferences are a distinct contributing factor to disciplinary referrals - and further, are distinct from behavior in the marshmallow task - highlights the importance of incorporating experimental measures of economic preferences into this literature. Children have agency and make decisions every day that are influenced by their own patience or willingness to take risks. Deciding whether or not to attend class, complete homework

[5]This would be consistent with findings from studies with adolescents (Castillo et al., 2011, 2019, 2018; Sutter et al., 2013)

assignments, or misbehave are choices that can have a compounding effect on human capital formation. Gaining a deeper understanding of the processes shaping human capital formation is essential to analyzing educational interventions that focus on each of these skills. In addition, data on economic preferences might be useful in identifying children who have the most to gain from such interventions.[6]

We also contribute to a growing literature in experimental economics on the impact of economic preferences on decision-making. This literature has reported on correlations between economic preferences, behaviors, and outcomes. Among adults, time discounting is associated with take-up of financial education (Meier and Sprenger, 2013), health outcomes (Khwaja et al., 2007; Harrison et al., 2010), and the distribution of wealth (Epper et al., 2020). Among adolescents, time discounting and risk preferences have been associated with greater likelihood of disciplinary referrals, school drop out, and adverse health behaviors (see List et al. (2018) for a review). Limitations of these studies include potential for omitted variable bias and reverse causality, since they most often consider a narrow set of skills and collect data contemporaneously.[7] None of the related work has considered the association of economic preferences measured as young as age 3-5 on later-life outcomes.

Our study augments this literature in several important ways. First, our comprehensive data allow us to disentangle the role of economic preferences from potentially confounding cognitive and executive function skills, which is typically not possible in related work. This limits the omitted variables problem. Second, we collect data on children at the earliest ages when experimental preference elicitation can reasonably be carried out. This is important since we seek to capture the impact of base economic

[6]This paper is thus along the lines of Almlund et al. (2011), who assert that personality and cognition are two important factors shaping IQ and later-life outcomes.

[7]Exceptions to the reverse causality problem include Castillo et al. (2019) who evaluate time preferences measured in pre-adolescence with high school dropout 5 years later, Alan and Ertac (2018) who examine a patience intervention on behavior one year later, Golsteyn et al. (2014) who examine hypothetical delayed consumption choices as a teenager on educational attainment as an adult and Epper et al. (2020) who show that earlier survey measures of patience relate to wealth distribution.

preferences. The related literature has shown rapid changes in risk, time, and social preferences during childhood - relative to younger children, older children tend to be more patient, less likely to take risks and more egalitarian (see Sutter et al. (2019) for a review). Third, our study separates the measured skills and field outcomes in time by multiple years, limiting the reverse causality problem. Finally, we take advantage of a field experiment with random assignment to evaluate the causal impact of early childhood programs on skills and outcomes.

Our study also has direct policy relevance. Our outcome measure captures misbehavior, which is shown to be a predictor of economic outcomes later in life, such as educational achievement and wages (Bowles et al., 2001; Heckman et al., 2006; Lang and Ruud, 1986) and high school drop-out rates (Alexander et al., 1997; Rumberger, 1995). Our study indicates that multiple early life skills are drivers for later-life disciplinary referrals and points to the value of early childhood interventions on this outcome. In addition, we study children in an urban, low-performing school district who may have the greatest amount to gain from early interventions relative to their higher-income counterparts.

In what follows, Section 2 describes the experimental design and discusses the construction of our variables. Section 3 presents our main results. Section 4 provides a discussion of the impact of the early childhood interventions on disciplinary referrals. Section 5 concludes.

2 Experimental Design and Variable Construction

2.1 Overview

Our data are generated from children who took part in the Chicago Heights Early Childhood Center (CHECC) project, a unique, large-scale field experiment implemented in a prototypical low-performing urban school district in Chicago Heights, Illinois (Fryer

et al., 2015). Nearly 80% of Chicago Heights is Black or Hispanic, and average per capita (pretax) income is $17,546, according to the 2020 US census. Households with children aged 3-4 were recruited to participate in CHECC between 2010 and 2014. Recruitment included a large marketing campaign within the Chicago Heights school district and in neighboring districts.[8] Participants were randomized to receive a preschool program, to receive a parenting program, or to a control group that did not receive any intervention for a period of 1-2 years, depending on year of entry and the child's age at entry. Our demographic and SES data come from surveys completed by parents.

We conducted assessments and experiments eliciting economic preferences with the CHECC children in 2010-11, 2012, and 2013. All sessions proceeded one-on-one with a trained staff member. Children from all CHECC programs were invited to participate, and did so either during school or during an evening or weekend time when parents were available to bring in their children.[9] Administrative data on disciplinary referrals was provided by the Chicago Heights school district after children started Kindergarten, and is therefore only available for the sub-set of CHECC children who ultimately attended Chicago Heights schools. Whether children attended Chicago Heights schools was determined by their primary residence at the start of Kindergarten. Table 1 provides a summary of the data available. Data on economic preferences were collected only once each year, while data on skills were collected both before and after CHECC interventions.

2.2 Childhood Skills

The skills assessment took 45 minutes to 1 hour to complete and always proceeded in the order described below.

Cognitive skills: We used standard, nationally-normed tests to measure cogni-

[8] All households in the area with children aged 3–4 years were eligible to participate, as long as the child did not have a learning difficulty that the program would be unable to address.

[9] During the evening/weekend activities, parents were usually compensated between $10 and $50 to bring in their children.

tive skills, including receptive vocabulary, letter and word recognition, writing, and math. Receptive vocabulary was tested using the Peabody Picture Vocabulary Test III (PPVT-III) (Dunn and Dunn, 1997), letter and word recognition was tested using the Woodcock-Johnson III Test of Achievement (WJ-III) letter-word sub-test, writing skills were tested using the WJ-III spelling sub-test, and math ability was tested using the WJ-III applied problems and quantitative concepts sub-tests (Woodcock et al., 2001). All of these tests are appropriate for children ages 2+ and start with easier questions that get progressively more difficult. For example, the spelling test first asks children to draw lines or trace letters and then progressively increases in difficulty to ask children to write letters and words. In our analysis, we use the percentile score for each test, which is obtained from the supplier based on their national norming sample.

Executive function skills: At the time we conducted these assessments, a normed standard for measuring executive function skills was not readily available. Therefore, we measured working memory, inhibitory control, emotion regulation, and attention in two ways. First, children participated in paper-and-pencil tasks developed by psychologists Clancy Blair and Michael Willoughby. In the working memory task, children were shown a series of images of animals in houses and then asked to recall which animals were in which houses (Blair and Willoughby, 2006a). In the inhibitory control task, children were asked to tap pictures of arrows corresponding to pictures they saw in a booklet, where the left arrow sometimes appeared on the right and the right arrow sometimes appeared on the left. Children had to inhibit their predominant response to tap the wrong arrow (Blair and Willoughby, 2006b).

Second, assessors blind to treatment completed a survey about the child's behavior at the end of the session. This survey - called the Preschool Self Regulation Assessment (PSRA) - measures emotion regulation and attention (Smith-Donald et al., 2007). In the survey, assessors chose the best fit statement in questions such as 'pays attention during instructions,' whereby each statement was scored, with more points indicating

higher emotion/attention. Since a normed standard for these tests are not available, we use a 'percent correct' or 'points attained out of maximum possible' score in our analysis.

2.3 Economic Preferences and Delay of Gratification

To measure economic preferences and delayed gratification, children received instructions for each experiment verbally and the monitor answered all questions prior to the child making choices (see Appendix 1). Props were used to aid understanding. Children were incentivized with candies or prizes. In different waves of data collection, the tasks differed slightly as described in Table 1. To address this potential issue, we standardized the measures for each preference, as described below. The economic preference assessment lasted about 20 minutes, and the delay of gratification task was typically conducted in a separate session and took 10-20 minutes to complete.

Time preferences: Children made a series of choices between fewer candies at the end of the same day or a greater number of candies at the end of the next day. To avoid impulsive decisions, all choices featured a front-end delay.[10] Table 1 describes the choices in each wave; for example, in 2010, children chose from the following: 4 today versus 5 tomorrow, 4 today versus 6 tomorrow, 4 today versus 7 tomorrow and 4 today versus 8 tomorrow. Children were informed that only one decision would be selected as the 'decision that counts' and would be paid at the end of the experiment.[11] Time preferences are represented in our analysis by the percentage of times the child was patient and chose to wait for the larger reward across all decisions.

Risk preferences: Children made a series of choices between 1 prize for sure and a larger number of prizes with some probability. As summarized in Table 1, in 2012 we varied the number of prizes in the risky box across decisions whereas in 2013 we

[10] In most waves, this meant the 'end of the day today' but in one instance we said 'the end of the activity today'. In practice, this meant the delay was between 1-5 hours for the same day.

[11] We implemented the decision that counts by storing selections from each round in small bins inside of an opaque bag. Children had to close their eyes and choose a bin at the end of the experiment.

varied the probability of getting prizes in the risky box across decisions. Children were informed that only one decision would be selected as the 'decision that counts' and would be paid at the end of the experiment.

A different risk elicitation was conducted in 2010-11 (Andreoni et al., 2019a). Children were shown 5 pencils in a jar. One pencil had a red mark on the bottom while others did not. Children chose how many pencils to take out of the jar. If none of the chosen pencils had a red mark, then children kept all of the pencils. If any of the pencils had a red mark on the bottom, then they forfeited the pencils. Since children could not see which pencil had a red mark until the end, choosing to take more pencils is associated with greater risk seeking preferences. In our analysis, risk preferences are represented by the percentage of times the child chose the safe option over the risky option across all decisions. For cohorts 2012 and 2013, this is equivalent to the proportion of safe decisions over 4 pairwise comparisons. For the cohort 2010, this is equivalent to the number of pencils not chosen out of 5.

Social preferences: Children participated in a dictator game in which each child was paired with another individual who did not participate in the experiments. In 2010-11, we followed a protocol common in the psychology field whereby children divided 5 candies between themselves and a fictitious character (a teddy bear). In 2012 and 2013, children divided 6 stickers between themselves and another anonymous child whereby we distributed the shared stickers to other children who did not participate. Social preferences are measured in the analysis by the percentage of stickers shared.

Delay of gratification: We complement the marshmallow task designed to measure delay of gratification, following the design in Mischel et al. (1972). Because this task was not the primary variable of interest for us, we only administered it to a subset of the children. In this task, children were given the option to wait for some amount of time and resist eating the fewer treats in front of them to receive more treats later. The marshmallow task is represented in the analysis by whether or not the child ate

the treat before 5 minutes had passed (since 5 minutes is the minimum wait time across years).

2.4 Disciplinary Referrals

Disciplinary data are available for grades K-6 between the 2011-12 and 2017-18 school years for children in the Chicago Heights school district. In the analysis, we first examine whether a child had any referral in any year and the number of referrals received. Next, we consider the severity of the referral. Disciplinary referrals can range from relatively minor problems, such as being late to class, to more serious offenses, such as striking another student. Teachers do not include a severity rating for the referral but they do include a description of the incident.

We therefore hired three independent coders (teachers from districts outside of Chicago Heights), who, blind to treatment, rated the severity of each incident on a scale of 1-7, where higher values signified lower severity of the referral. We asked coders to rate each incident based on the following, 'In your view, how badly did the student behave this week?' from 1 (Extremely badly) to 7 (Not badly at all). These values were then reversed in sign (-1 to -7) for ease of interpretation in the analysis.[12]

Our main outcome variable is misbehavior as measured by weekly disciplinary referrals. We construct three variables, 1) the propensity to receive at least one referral in a school year, 2) the number of referrals received in a school year and 3) the severity of the referrals received in a school year. The last variable is based on our independent coder ratings.[13]

[12]Since some teachers report referrals on a daily basis (e.g., 'student was late to class today') and others report referrals on a weekly basis (e.g., 'student was late to class 3 times this week'), for the ratings and number of referrals, we compiled a dataset that contained the disciplinary referrals that occurred over the course of a full week. Thus, all referrals over the course of a week are aggregated for counts and rating.

[13]In this analysis, children without a disciplinary referral are coded as having a rating of 8.

3 Results

3.1 Overview

We gathered data on skills and economic preferences for 856 unique children, but only 376 of these children attended one of the 9 schools in the Chicago Heights district and therefore have data on disciplinary referrals on file.[14] In sub-section 3.2, we use the full sample of children with skills data to discuss the interrelationships between skills in early childhood. In sub-section 3.3, we use the smaller sample of Chicago Heights children who also have disciplinary referral data to discuss the association between early childhood skills and disciplinary referrals later in life. Appendix 2 summarizes the sample selection. We show in Appendix 3, Table A1 that children with skills and preference data do not differ from those without available data with respect to prevalence of disciplinary referrals. This makes sense since the missing data arises from the location of the child's house, and the neighborhoods in and out of district are adjacent. Yet, there are some differences between samples along observables.

Similar to the demographics of Chicago Heights, IL, our sample is diverse and relatively low income. Within the larger sample, 40% are Hispanic, 37% are Black and the remainder are White. The average age of the child's mother at the birth of her first child is about 23 years old, and about 76% of mothers have at least a high school degree. Average household income is $28,875. Children were on average 3.8 years old when they completed the skills and economic preference assessments.

Within the Chicago Heights sub-sample, 59% are Hispanic, 21% are Black and the remainder are White. The age of the child's mother at the birth of her first child is about 21 years old, and about 60% of mothers have at least a high school degree. Average household income is $20,910. Children were on average 3.8 years old when

[14] This smaller sub-sample was expected since the CHECC program intentionally drew on children from many different districts, including the home district of Chicago Heights, and even provided free bus service for some students outside the district.

they completed the skills and economic preference assessments. Appendix 3, Table A2 summarizes the characteristics of the full and the sub-sample. There are a few notable differences - for example, the suburb of Chicago Heights has more Hispanic and fewer Black households than the surrounding communities, and this is reflected in the racial and ethnic composition of our samples. The larger sample is also somewhat more educated and has a higher income than the Chicago Heights sample.

Table 2 provides a summary of the children's skills and economic preferences. Children score on average in the 30^{th} to 46^{th} percentile on cognitive skills. This is somewhat below the national average. Average scores on the executive functioning tests and assessor reports are between 34% and 73%. Appendix 3, Table A3 provides the same summary for the subsample with disciplinary data.

Children choose the patient option in the time preference elicitation in 33% of the cases, and the safe option in the risk preference elicitation 44% of the time. Children share on average 34% of their endowment in the social preference elicitation. By comparison, research with adults show sharing rates of about 20% (Forsythe et al., 1994). Finally, 43% of children eat the marshmallow within five minutes. The children in our sample are more willing to wait than those in the original Mischel study, where 75% of the children ate the marshmallow (Mischel et al., 1972). Part of the difference could be attributed to the fact that we chose a shorter length of time to allow children to wait. Appendix 3, Figure A2 provides the economic preference data by gender and race. There, we see patterns consistent with related work: for example, we observe that girls are more patient and more risk averse than boys (Sutter et al., 2019).[15]

Figure 1 shows the percentage of children in our sample who received at least one disciplinary referral by grade in school. Disciplinary referrals are more prevalent in

[15]We also see that the Black children in our sample are more impatient than White or Hispanic children in the time preference task and that White children are less likely to eat the marshmallow in the delay task, consistent with findings by Andreoni et al. (2019b) and Castillo et al. (2011). Interestingly, using adolescent data from the Chicagoland area, Cotton et al. (2020) find similar time preference data over leisure preferences.

later elementary school grades than in early years. Kindergarten children are the least likely to receive a referral (8% received a referral), and children in 5th grade are most likely to receive a referral (23% received a referral). Most disciplinary referrals are concentrated among a small percentage of children. Of the children in our sample, 62% have never received a referral. The average number of referrals in a year is 0.47 (s.d. 1.94), and 17% of the children received above the average number of referrals.

3.2 Interrelationships Between Skills and Preferences

We next evaluate the interrelationships between early life skills and economic preferences. Table 3 provides correlations between skills and preferences. We observe that while measures of cognitive skills and executive functions are correlated with one another, they are importantly much less correlated with economic preferences. Interestingly, while this correlation has been observed to be quite strong in adults (Burks et al., 2009; Dohmen et al., 2010), the correlation amongst our 3-5 year-old sample is much weaker. Also, we find that economic preferences are not strongly correlated with one another, providing some optimism that such variables can provide unique information for predicting later life outcomes.

We next conduct exploratory factor analysis with our full sample of children to evaluate whether the 13 measures we collected can be mapped to a smaller sub-set of latent factors. This is useful for understanding interrelationships of early life skills. It is also useful because we can then use an index of latent factors to better understand the associations of early life skills to outcomes. This can be more powerful and informative than evaluating the association of each measure individually.

Table A4 in Appendix 3 summarizes the factor loadings from our exploratory analysis after using oblique rotation, which allows for the natural assumption that factors can be correlated. Our results suggest four latent factors are represented by our 13

measures.[16]

As we may have predicted, our measures of cognitive skills and executive functions each map on to two different factors in direct alignment with our earlier categorization of these measures. More surprisingly, we find that economic preferences are distinct from cognitive skills and executive functions. They map on to two separate factors, with time and social preferences loading together and risk preferences alone. Further, the delay of gratification task does not map to any of the four factors. We observe that the two factors for economic preferences have low eigenvalues, providing weak evidence of association. In the subsequent section, we use a factor for cognitive skills, a factor for executive functions, and examine each economic preference separately.[17]

3.3 Associations with Disciplinary Referrals

We next examine the correlations of economic preferences, cognitive skills, and executive functions measured at age 3-5 years old with the propensity to receive disciplinary referrals in elementary school. For the main analysis, we use an index for cognitive skills and executive functions. We include economic preferences and the delay of gratification measure separately.

Our main outcome variables are the probability of having at least one referral in a given school year and the average severity of the misbehavior reported in the referrals for that year. As aforementioned, the severity of the referrals was compiled by independent raters using a Likert scale of 1-7 from (1) very severe (e.g., child behaved extremely badly) to (7) not severe (e.g., not badly at all). These values were then reversed in sign for ease of interpretation.

Table 4 reports the empirical results. We include each year for which we observe

[16] We estimate the latent factors using Bartlett's method (Bartlett, 1937). This method produces unbiased factors that can be used in linear regressions to produce consistent estimates (Skrondal and Laake, 2001). Horn's parallel analysis (Horn, 1965) is used to determine the number of factors to keep from the exploratory factor analysis. This yields the following adjusted eigenvalues for four factors: 3.340, 1.605, 1.014 and 1.005.

[17] Analysis using factors for economic preferences and other averaging is in Appendix 3.

the child in elementary school as an observation, and all analyses control for the child's demographic characteristics and cluster standard errors at the child level. The first two columns in Table 4 report on the propensity to receive a referral, and the last two columns report on the severity. Each column in the table controls for each of the economic preference measures, the cognitive skill index, the executive function index and demographics.[18]

We find that across all empirical specifications, higher executive functions in early childhood are associated with a reduction in misbehavior in elementary school. This is true both for the propensity of having a referral and the severity of referrals. For example, a one standard deviation increase in the executive function index results in a 1.9 percentage point decline in the probability of receiving a referral. The baseline likelihood of disciplinary referrals is 14.7%. By contrast, cognitive skills in early childhood are not associated with changes in the propensity to have disciplinary referrals or the severity of the referral.

Most interestingly, early life economic preferences - specifically time preferences - are separately associated with prevalence and severity of disciplinary referrals. Children who are willing to wait for larger, later rewards at ages 3-5 are less likely to receive disciplinary referrals in elementary school. For example, a one standard deviation increase in the willingness to wait is associated with a reduction of the probability of receiving a referral by 2.3 percentage points. However, risk preferences, social preferences and delay of gratification as measured by the marshmallow task are not strong predictors of disciplinary referrals later in life.

Finally, we find strong differences by gender and race in our sample: Boys and Black children are significantly more likely to receive disciplinary referrals. Boys are 10% more likely to receive a disciplinary referral compared to girls, and Black children

[18]We use two specifications in the analysis. First, we regress disciplinary referrals on all variables on the sub-sample of children who have data on all measures. Second, to increase precision in our estimates, we replace missing item responses by the mean of the variable and include dummies that equal 1 if the item response is missing.

are 8% more likely to receive a disciplinary referral than their racial counterparts. This finding is aligned with related work in economics (e.g. Castillo et al., 2011; Bertrand and Pan, 2013), which shows that boys and Blacks are more likely to have disciplinary referrals. More broadly, a large literature has documented a 'discipline gap' between Blacks and non-Blacks, such that Blacks receive more referrals (Gregory et al., 2010).

Our results are robust to a variety of alternative specifications, including using number of disciplinary referrals as an outcome (see Appendix 3, Table A5). They are also robust to using two factors for economic preferences, averaging across economic preferences or adding additional family and household characteristics (see Appendix 3, Tables A6-A10). Using each cognitive and executive function skill separately yields similar results (see Appendix 3, Table A11). The executive function assessment with the greatest explanatory power is the self-regulation assessment on attention. That is, a child's ability to pay attention and patience in early childhood are both predictors of later-life outcomes.

4 Mechanisms of mis-behavior

We next leverage the randomization inherent in CHECC to evaluate whether there is an impact of early childhood programs on disciplinary referrals, and to understand the role of skills and economic preferences as moderators or mediators. Our analysis exploits the fact that the CHECC treatments aimed to boost the cognitive and executive functioning skills of participating children, and we have data on these skills both before and after the intervention. Upon signing up for CHECC, households were randomly assigned to one of the following treatments:

- **Preschool (PK):** Children attended a free, full-day preschool for 1-2 years. Children were randomized to one of three curricula: a program focused on literacy and math, a program focused on executive functions, and a program focused on

both skills with a parent involvement component.

- **Parent Academy (PA):** Parents attended a bi-monthly, 2-hour class to learn how to teach to their children. Parents received incentives of up to $3,000 based on attendance, homework completion, and their child's performance. Parents were randomized to receive part of those incentives as cash or as a deposit in the child's post-secondary education account.

- **Control group:** This group did not receive educational interventions from us.[19]

Fryer et al. (2020) report that the Preschool program improved cognitive skills, while Fryer et al. (2015) find that the Parent Academy program improved executive function skills. Related papers find that both programs also had an impact on fairness views (Cappelen et al., 2020), but no impact on time preferences (Andreoni et al., 2019b).

Table 5 summarizes the impact of the CHECC program on skills for the sub-sample of children who also have disciplinary referral data. Similar to (Fryer et al., 2020), we find an impact of Preschool on cognitive skills in this sub-sample, but the impact of Parent Academy on executive functioning skills is limited, potentially due to low power given the smaller sample size relative to Fryer et al. (2015). We also do not find an impact on time preferences, which is similar to the findings reported in the larger sample of (Andreoni et al., 2019b).

We next turn to evaluating whether CHECC had an impact on disciplinary referrals. There is a general consensus in the literature that the effects of preschool programs

[19] Another group was randomized to a shorter, summer preschool program called Kinderprep. We do not evaluate that program here due to its smaller sample size and due to the fact that it is in effect a lower-touch variant of the Preschool program. The total number of children randomized to each program over the four years was Preschool, excluding Kinderprep N=532; Parent Academy, N=316; Control group, N=1,132. We have disciplinary referral data for all children attending Chicago Heights schools. We use disciplinary referral data on 501 children for the treatment effect analysis i.e., Preschool excluding Kinderprep, N=187 (35.15 %); Parent Academy, N=129 (40.82%) and Control group, N=185 (16.34%). The smaller sub-sample is one limitation of this analysis. The smaller proportion of control group present in Chicago Heights is due to the initial randomization, which over-sampled Chicago Heights residents for the treatments in some years.

'fade out' in the years following the program (Duncan and Magnuson, 2013). Since disciplinary referrals are collected through sixth grade, following this literature, we might expect any effects to be muted.

Table 6 presents the impact of CHECC on disciplinary referrals, using both whether the child had any referral in a year and the severity of the referral as outcome measures.[20] Table 6 shows that random assignment to the Parent Academy program had a statistically significant effect of reducing the likelihood of receiving any disciplinary referral by 27%. We also find suggestive evidence that random assignment to Preschool reduced the likelihood of having any disciplinary referral by 15%, but this result is not statistically significant at conventional levels. We also find that, consonant with our earlier analysis, baseline skills are predictive of disciplinary referrals, but the coefficients are not statistically significant at conventional levels, likely at least in part due to the smaller sample size.[21]

In Columns (2)-(3) and (5)-(6) of Table 6, we include controls for the cognitive and executive functioning skills at the completion of the intervention, respectively. The executive function skills at the end of the program are strongly associated with disciplinary referrals, whereas cognitive skills are not correlated. Adding these variables does not change the treatment effects. Columns (4)-(6) replicate this analysis for severity of referrals and show similar results.

Taken together, the above results suggest that the CHECC treatments do not have a robust effect on the mediating variables of cognitive skills and executive functioning in the sub-sample that we study, but that at least one of the treatments does have an impact on reducing disciplinary referrals. This finding suggests that the effect of Parent Academy on discipline is direct and not mediated by cognitive skills and

[20] Appendix Table A12 reports the same result using the number of weekly referrals in a year and Appendix Table A14 reports the same result disaggregating the program curricula.

[21] Note that the sample used in this analysis differs slightly from that used for Table 4, see Appendix Figure A1. Table 4 examines the effect of overall time preferences of each child on disciplinary referrals while Table 6 controls for the effect of only baseline time preferences i.e. time preferences collected pre-treatment.

executive functions.

As a robustness test, we account for the probability that children for whom we have disciplinary data might be different from the children for whom we do not by estimating inverse probability weighted (IPW) treatment effects in Table A17. IPWs are estimated from running a probit regression on a binary variable indicating whether disciplinary referral data is available for the child using gender, race, and age at assessment as controls. Empirical results remain similar upon weighing the coefficients by IPWs.[22]

Importantly, the design of the CHECC treatments affords an evaluation of alternative policy options. The Parent Academy treatment was designed to enhance parenting practices, and in turn children's available choice sets, rather than to boost cognitive and executive function skills directly. Table 6 demonstrates that interventions targeted at parenting practices can improve child outcomes even when they do not affect skills or preferences. While it is still early to know the long-term outcomes of the interventions, the evidence so far suggests they might be durable.[23]

Finally, one might wonder about the generalizability of our insights. Beyond our exploration of the results in this sample, how far can these results expand to other populations of people and situations? To lend some insights, we follow the List (2020) SANS conditions in our reporting.[24] First, in terms of selection, our sample includes students from the CHECC program, which were a representative group of pre-K students from Chicago Heights and the surrounding area. Importantly, Chicago Heights has observables similar to other inner-city school districts around the US. In terms of attrition, our compliance rates with treatment were about 80% (Fryer et al., 2015,

[22]We have a sub-sample of children included in the treatment analysis for whom we have time preference data collected after completion of the intervention. Due to the very small sub-sample size, we relegate the table adding endline time preferences in Table 6 to the Appendix, Table A15. In this sub-sample we do not find a statistically significant effect of the Parent Academy program. Thus, we cannot distinguish if time preferences mediate the effect of Parent Academy.

[23]Jackson (2018) shows that teachers can affect students' behavior and this can have a separate effect on academic performance. That analysis, however, cannot distinguish whether teachers affect behaviors by altering choice sets or by changing attitudes. Our results show that both channels are possible.

[24]SANS abbreviates Selection, Attrition, Naturalness and Scaling.

2020), while our administrative data on disciplinary referrals captures all children who attended the Chicago Heights school district.

Considering naturalness of the choice task, setting, and time frame, we use a combination of standardized tests and artefactual/framed field experimental manipulations. Thus, our setting is one in which subjects are naturally engaged, and includes a mixture of natural choices and some choices that place students on an artificial margin in that they would not typically make such choices. In terms of scaling, CHECC was designed as an early childhood program based on scaling, thus careful consideration of the non-negotiables of CHECC, including most importantly human capital, was of first order importance. Finally, since our key results are WAVE1 insights (List, 2020), replications need to be completed to understand if the estimates manifest in other school districts.

5 Conclusion

We use experiments to collect data on a wide range of skills and economic preferences with children when they are ages 3-5 years old, and we return to the same children when they are ages 5-12 years old to collect administrative data on disciplinary referrals, a key predictor of adult outcomes. We find that cognitive and executive functioning skills map to two distinct latent factors, and that economic preferences map to separate latent factors. We also find that executive functions and time preferences are associated with the propensity for disciplinary referrals several years later.

We use disciplinary referrals in middle-childhood as our outcome measure, which has been previously linked to adult outcomes such as educational attainment and labor market earnings (e.g. Segal, 2013). Finding that early life executive functions and time preferences predict disciplinary referrals in middle-childhood points to the importance of considering such skills when designing early childhood assessments and optimal education policy. While the literature by James Heckman and co-authors (Cunha et al.,

2010; Almlund et al., 2011) has often discussed the role of cognitive skills captured by academic achievement tests and 'soft skills' such as character and temperament, executive functions are higher-order cognitive skills that have been discussed less frequently in the economics literature yet are of importance in the education production function. An exception is recent work by Heckman et al. (2020) that explores the impact of a home visiting program in China on a comprehensive set of skills, including on executive functions (which the authors refer to as a component of 'skills for learning.') Importantly, time preferences are prominent in the literature on determinants of adult behavior (Barsky et al., 1997; Meier and Sprenger, 2013; Cohen et al., 2020). Evaluations of how these preferences map to behavior in childhood are much more limited. We show that such skills are relevant even among very young children, which suggests that such skills should be routinely evaluated when considering the determinants of child behavior and outcomes.

Early studies of the 'marshmallow test' pointed to the importance of delay of gratification in shaping human behavior. The interpretation that this task captured delay of gratification and not other aspects of early childhood development or the child's environment has come under some scrutiny recently (Kidd et al., 2013). Our work collects a comprehensive set of measures and suggests that while the marshmallow test alone cannot predict behavior, other related constructs of executive functioning and time preferences are separately predictive. Executive functions are well understood by psychologists as a key developmental construct.

Time preferences collected by standard economic experiments differ from either the inhibitory control measure within executive function assessments or the marshmallow test. Time preferences capture a deliberate decision (because all rewards are associated with a front-end delay), while the others have an element of impulse control. Our novel and perhaps surprising finding is that the deliberative decisions of toddlers and preschoolers are informative of outcomes for these children years later. Economists

studying human capital formation should consider economic preferences as a unique driver of behavior.

Our finding that such time preferences are also predictive of behavior among young children provides further evidence of the importance of this skill as a separate predictor, as has been documented with adults. As we point out in List et al. (2018), understanding the drivers of behavior at young ages is important because these help us understand children and, by understanding children, we can better understand adults.

Another important finding in our study is that early childhood interventions can affect misbehavior in middle childhood, but that these effects appear not to be mediated by the skills that we measure or by economic preferences. This suggests that early childhood interventions operate directly or through alternative mediators that we did not consider - e.g., other 'soft skills,' such as personality and character, or by changing parents' behavior. This is an area for further research.

Interventions aimed at encouraging children to invest in human capital formation should consider time preferences as an additional driver of behavior. For example, children who discount the future most at early ages might be at greater risk for poor outcomes in adulthood. This points to the importance of identifying children in need of additional scaffolding and improving our understanding of the factors affecting children's preferences.

References

Alan, S. and Ertac, S. (2018). Fostering patience in the classroom: Results from randomized educational intervention. *Journal of Political Economy*, 126(5):1865–1911.

Alexander, K. L., Entwisle, D. R., and Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, pages 87–107.

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education*, volume 4, pages 1–181. Elsevier.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Andreoni, J., Di Girolamo, A., List, J. A., Mackevicius, C., and Samek, A. (2019a). Risk preferences of children and adolescents in relation to gender, cognitive skills, soft skills, and executive functions. *Journal of Economic Behavior & Organization*.

Andreoni, J., Kuhn, M. A., List, J. A., Samek, A., Sokal, K., and Sprenger, C. (2019b). Toward an understanding of the development of time preferences: Evidence from field experiments. *Journal of Public Economics*, 177:104039.

Barsky, R., Juster, F., Kimball, M., and Shapiro, M. (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *Quarterly Journal of Economics*, 112(2):537–579.

Bartlett, M. (1937). The statistical conception of mental factors. *British Journal of Psychology*, 28:97–104.

Benjamin, D., Laibson, D., Mischel, W., Peake, P., Shoda, Y., Wellsjo Steiny, A., and Wilson, N. L. (2020). Predicting mid-life capital formation with pre-school delay

of gratification and life- course measures of self-regulation. *Journal of Economic Behavior and Organization*.

Benjamin, D. J., Brown, S. A., and Shapiro, J. M. (2013). Who is 'behavioral'? cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11(6):1231–1255.

Bertrand, M. and Pan, J. (2013). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal - Applied Economics*, 5(1):32–64.

Bettinger, E. and Slonim, R. (2007). Patience among children. *Journal of Public Economics*, 91(1):343–363.

Blair, C. and Willoughby, M. (2006a). Measuring executive function in young children: Operation span. *Chapel Hill, NC: The Pennsylvania State University and The University of North Carolina at Chapel Hill*.

Blair, C. and Willoughby, M. (2006b). Measuring executive function in young children: Spatial conflict ii: Arrows. *Chapel Hill, NC: The Pennsylvania State University and The University of North Carolina at Chapel Hill*.

Bowles, S., Gintis, H., and Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, 39(4):1137–1176.

Burks, S. V., Carpenter, J. P., Goette, L., and Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19):7745–7750.

Cappelen, A. W., List, J. A., Samek, A., and Tungodden, B. (2020). The effect of early-childhood education on social preferences. *Journal of Political Economy*, 128(7):2739–2758.

Carneiro, P. M. and Heckman, J. J. (2003). Human capital policy. In *Inequality in America: What role for human capital policies?* Cambridge, MA: MIT Press.

Castillo, M., Ferraro, P. J., Jordan, J. L., and Petrie, R. (2011). The today and tomorrow of kids: Time preferences and educational outcomes of children. *Journal of Public Economics*, 95(11):1377–1385.

Castillo, M., Jordan, J. L., and Petrie, R. (2018). Children's rationality, risk attitudes, and field behavior. *European Economic Review*, 102:62–81.

Castillo, M., Jordan, J. L., and Petrie, R. (2019). Discount rates of children and high school graduation. *The Economic Journal*, 129(619):1153–1181.

Cohen, J., Ericson, K. M., Laibson, D., and White, J. M. (2020). Measuring Time Preferences. *Journal of Economic Literature*, 58(2):299–347.

Cotton, C., Hickman, B. R., List, J. A., Price, J. P., and Roy, S. (2020). Productivity versus motivation in adolescent human capital production: Evidence from a structurally-motivated field experiment. National Bureau of Economic Research, Working Paper No. 27995.

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3):883–931.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives*, 15(2):213–238.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100:1238–1260.

Duncan, G. J. and Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, 27(2):109–32.

Dunn, L. M. and Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. American Guidance Service.

Epper, T., Fehr, E., Fehr-Duda, H., Kreiner, C. T., Lassen, D. D., Leth-Petersen, S., and Rasmussen, G. N. (2020). Time Discounting and Wealth Inequality. *American Economic Review*, 110(4):1177–1205.

Forsythe, R., Horowitz, J., Savin, N., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3):347–369.

Fryer, R. G., Levitt, S. D., List, J. A., and Samek, A. (2020). Introducing CogX: A new preschool education program combining parent and child interventions. National Bureau of Economic Research, Working Paper No. 27913.

Fryer, R. J. G., Levitt, S. D., and List, J. A. (2015). Parental incentives and early childhood achievement: A field experiment in Chicago Heights. National Bureau of Economic Research, Working Paper No. 21477.

Golsteyn, B. H. H., Gronqvist, H., and Lindahl, L. (2014). Adolescent Time Preferences Predict Lifetime Outcomes. *Economic Journal*, 124(580, F):F739–F761.

Gregory, A., Skiba, R. J., and Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, 39(1):59–68.

Harrison, G. W., Lau, M. I., and Rutström, E. E. (2010). Individual discount rates and smoking: Evidence from a field experiment in Denmark. *Journal of Health Economics*, 29(5):708–717.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–2086.

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, 54(1):3–56.

Heckman, J. J., Liu, B., Lu, M., Zhou, J., et al. (2020). Treatment effects and the measurement of skills in a prototypical home visiting program. Technical report, Institute of Labor Economics (IZA).

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrik*, 30(2):179–185.

Khwaja, A., Silverman, D., and Sloan, F. (2007). Time preference, time discounting, and smoking decisions. *Journal of Health Economics*, 26(5):927–949.

Kidd, C., Palmeri, H., and Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, 126(1):109–114.

Lang, K. and Ruud, P. A. (1986). Returns to schooling, implicit discount rates and black-white wage differentials. *The Review of Economics and Statistics*, pages 41–47.

List, J., Petrie, R., and Samek, A. (2018). How experiments with children can inform economics. *Work in Progress*.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. *National Bureau of Economic Research - Working Paper No. 27535*.

Meier, S. and Sprenger, C. D. (2013). Discounting financial literacy: Time preferences and participation in financial education programs. *Journal of Economic Behavior & Organization*, 95:159–174.

Mischel, W., Ebbesen, E. B., and Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, 21(2):204.

Mischel, W., Shoda, Y., and Rodriguez, M. I. (1989). Delay of gratification in children. *Science*, 244(4907):933–938.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32(3):583–625.

Schlam, T. R., Wilson, N. L., Shoda, Y., Mischel, W., and Ayduk, O. (2013). Preschoolers' delay of gratification predicts their body mass 30 years later. *The Journal of Pediatrics*, 162(1):90–93.

Segal, C. (2013). Misbehavior, education, and labor market outcomes. *Journal of the European Economic Association*, 11(4):743–779.

Shoda, Y., Mischel, W., and Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology*, 26(6):978.

Skrondal, A. and Laake, P. (2001). Regression among factor scores. *Psychometrica*, 66(4):563–576.

Smith-Donald, R., Raver, C. C., Hayes, T., and Richardson, B. (2007). Preliminary construct and concurrent validity of the preschool self-regulation assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, 22(2):173–187.

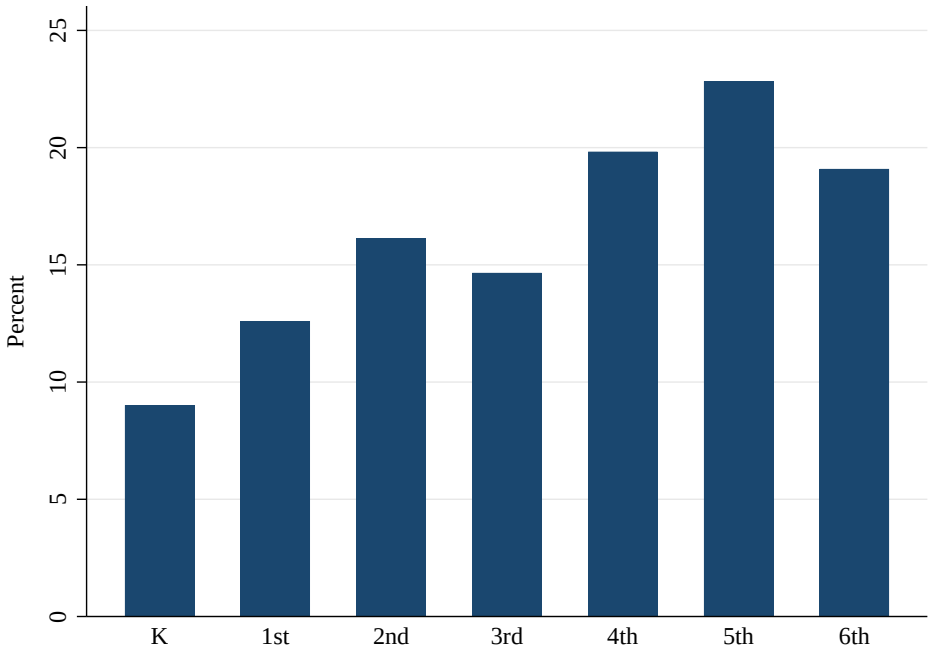
Sutter, M., Kocher, M. G., Glätzle-Rützler, D., and Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *The American Economic Review*, 103(1):510–531.

Sutter, M., Zoller, C., and Glätzle-Rützler, D. (2019). Economic behavior of children and adolescents–a first survey of experimental economics results. *European Economic Review*, 111:98–121.

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing.

Zelazo, P. D., Blair, C. B., and Willoughby, M. T. (2016). Executive function: Implications for education. NCER 2017-2000. *National Center for Education Research*.

Figure 1: Proportion of Children with Disciplinary Referrals by Grade



Note: This figure shows the percentage of children who had at least one disciplinary referral in each school year

Table 1: Summary of Data

Year	Skills	Preferences			Delay	Discipline
		Time	Risk	Social		
2010-11	Yes	4v5, 4v6, 4v7, 4v8.	Pencil task	5 candies, teddy.	Marshmallow, 8 min.	K
2012	Yes	3v3, 2v3, 1v3.	1 v 50% chance of 5, 4, 3, 2.	6 stickers, child.	Marshmallow, 5 min.	K-1st
2013	Yes	2v3, 2v4, 2v5, 2v6.	1 v 20%, 25%, 33%, 50% chance of 2.	6 stickers, child.	Candy, 15 min.	K-2nd
2014	Yes					1st-3rd
2015						2nd-4th
2016						3rd-5th
2017						4th-6th

Note: Time data list decisions and trade-offs, i.e. "4v5" means "choose 4 candies today or 5 candies tomorrow." Risk data list decisions, i.e. "1 candy for sure or 50% chance of 5 candies." Social data list decisions, i.e. "Divide 5 candies between yourself and teddy. The 'Pencil task' is number of pencils taken out of jar. Child keeps all pencils if no red mark on a pencil, o/w gets no pencils. The first wave of preference data collection spanned two calendar years, while the remaining waves occurred within one year. The years for which discipline data is available are academic years 2011-12 through 2017-18.

Table 2: Summary Statistics

Cognitive Skills	Percentile Score	SD
Letter and Word Recognition (WJ-III)	41.94	29.46
Writing Skills (WJ-III)	36.41	28.94
Math - Applied Problems (WJ-III)	40.18	26.45
Math - Quantitative Concepts (WJ-III)	46.43	23.98
Receptive Vocabulary (PPVT-III)	30.45	24.93
Executive Functions	Percent Correct	SD
Attention (PSRA Assessment)	73.02	25.76
Emotion Regulation (PSRA Assessment)	59.28	24.04
Working Memory (Blair & Willoughby)	34.10	31.52
Inhibitory Control (Blair & Willoughby)	66.30	25.82
Economic Preferences	Decisions	SD
Time Preferences (% Patient Decisions)	32.91	33.69
Risk Preferences (% Safe Decisions)	43.57	34.14
Social Preferences (% Shared)	33.62	26.47
Delay of Gratification (% Ate Marshmallow)	43.00	48.39

Note: Cognitive measures are reported as percentile of age distribution. Executive function measures are reported as percent correct relative to total possible score. Cognitive skills are based on the first test administered to the child, and economic preferences are an average in the case the child completed the task more than once. This table reports the data for the full sample (N=854). Similar measures are obtained with the sub-sample for whom we have disciplinary data (N=376), see Appendix Table A3.

Table 3: Cross-Correlation Table

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1):Letters	1.00												
(2):Writing	0.45***	1.00											
(3):Math Problems	0.44***	0.43***	1.00										
(4):Math Concepts	0.48***	0.36***	0.53***	1.00									
(5):Vocabulary	0.28***	0.29***	0.49***	0.45***	1.00								
(6):Attention	0.18***	0.32***	0.36***	0.18***	0.23***	1.00							
(7):Emotion	0.18***	0.29***	0.29***	0.20***	0.25***	0.52***	1.00						
(8):Working Memory	0.10***	0.22***	0.28***	0.02	0.20***	0.44***	0.29***	1.00					
(9):Inhib Control	0.14***	0.26***	0.36***	0.07**	0.26***	0.52***	0.27***	0.47***	1.00				
(10):Time	-0.11***	-0.05	-0.06	0.03	-0.01	-0.02	-0.08**	-0.08**	-0.09**	1.00			
(11):Risk	0.01	0.05	0.00	-0.06	-0.04	0.08**	0.07*	0.06	0.02	-0.03	1.00		
(12):Social	0.07*	0.05	0.00	0.06	0.05	0.02	-0.00	-0.01	-0.03	0.06*	0.02	1.00	
(13):Marshmallow	-0.02	-0.04	-0.08**	0.03	-0.06*	-0.13	-0.09**	-0.10***	-0.15***	0.00	-0.01	0.02	1.00

Notes: Cognitive measures are reported as percentile of age distribution. Executive function measures are reported as percent correct relative to total possible score. Cognitive skills are based on the first test administered to the child, and economic preferences are an average in the case the child completed the task more than once.Correlations are based on full sample of children who have skill and preference measures (740)

Table 4: Propensity to Have Disciplinary Referrals and Severity of Referrals - Probit Marginal Effects Regression and Ordered Logit Regression

	(1) 0-1	(2) 0-1	(3) Severity	(4) Severity
Percent of patient decisions [0,1]	-0.07*	-0.07**	-0.61*	-0.62*
	(0.04)	(0.03)	(0.36)	(0.33)
Percent of safe decisions [0,1]	-0.04	-0.04	-0.36	-0.42
	(0.04)	(0.03)	(0.36)	(0.32)
Percent shared [0,1]	0.01	0.00	0.10	0.06
	(0.04)	(0.04)	(0.42)	(0.37)
Ate marshmallow before 5min	0.01	0.00	0.15	0.02
	(0.02)	(0.02)	(0.22)	(0.22)
Executive function index	-0.02*	-0.02**	-0.20*	-0.21**
	(0.01)	(0.01)	(0.10)	(0.09)
Cognitive index	-0.01	-0.01	-0.10	-0.09
	(0.01)	(0.01)	(0.15)	(0.13)
Male (d)	0.11***	0.11***	1.05***	1.00***
	(0.02)	(0.02)	(0.22)	(0.20)
Age at experiment	0.00	0.00	0.03**	0.03**
	(0.00)	(0.00)	(0.02)	(0.01)
Black (d)	0.09*	0.12**	1.01***	1.14***
	(0.05)	(0.05)	(0.35)	(0.31)
Hispanic (d)	-0.01	0.01	0.17	0.31
	(0.03)	(0.03)	(0.31)	(0.26)
Observations	1265	1554	1313	1616
R2	0.14	0.13	0.10	0.09
Number of children	313	369	318	375

Note: Individual level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment not shown. Columns (2) and (4) replace missing experimental data with mean values and add dummies for each experimental measure missing. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010

Table 5: Average Treatment Effects on Skills

	(1) Cognitive Index	(2) Executive Function Index	(3) % Patient Decisions
PA	0.12	0.02	-0.14
	(0.12)	(0.16)	(0.33)
PK	0.32***	0.20	0.05
	(0.11)	(0.15)	(0.29)
R2	0.78	0.67	0.25
N	376	376	135

Notes: This table reports treatment effects on skills. Column 1 uses cognitive score at the end of the treatment year as the outcome (calculated as the mean of the standardized values of each of the subtests). Column 2 uses Executive Function score at the end of the treatment year as the outcome, and Column 3 uses proportion of patient decisions post-treatment as the outcome. Regressions control for cognitive baseline scores and Executive Function baseline scores, gender, race, home language, age at test date, matched pair grouping, test form, number of previous assessments, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth, birthweight, assessor and an indicator if the pretest is from a previous year. * $p<0.10$, ** $p<0.05$, *** $p<0.010$

Table 6: Average Treatment Effects on Propensity to Have Disciplinary Referrals and Severity of Referrals

	(1) 0-1	(2) 0-1	(3) 0-1	(4) Severity	(5) Severity	(6) Severity
PK	-0.15	-0.10	-0.17	-0.23	-0.16	-0.28
	(0.12)	(0.13)	(0.13)	(0.22)	(0.23)	(0.23)
PA	-0.27**	-0.25**	-0.33***	-0.46**	-0.42*	-0.56**
	(0.13)	(0.12)	(0.13)	(0.22)	(0.22)	(0.23)
Cognitive Baseline	-0.05	-0.05	-0.09	-0.10	-0.11	-0.15
	(0.08)	(0.08)	(0.08)	(0.15)	(0.16)	(0.16)
Executive Function Baseline	-0.10	-0.10	-0.04	-0.16	-0.16	-0.06
	(0.06)	(0.06)	(0.07)	(0.12)	(0.12)	(0.12)
Baseline Time Preferences	-0.07	-0.09	-0.12	-0.07	-0.09	-0.16
	(0.09)	(0.09)	(0.09)	(0.17)	(0.17)	(0.17)
Cognitive End Year		0.00	0.12		0.01	0.19
		(0.09)	(0.09)		(0.15)	(0.15)
Executive Function End Year			-0.29***			-0.47***
			(0.08)			(0.13)
Observations	2264	2264	2264	2313	2313	2313
R2	0.166	0.169	0.185	0.104	0.105	0.113
Number of children	486	486	486	501	501	501

Notes: This table reports treatment effects on disciplinary behavior. Column 1, 2 and 3 use a binary variable as the outcome indicating whether the child has had at least one disciplinary referral. Columns 4, 5 and 6 use the severity (rated by independent coders) of the disciplinary referral as the outcome variable. All regressions control for cognitive baseline scores and Executive Function baseline scores, baseline time preferences (with missing values replaced as 0 and including a dummy indicating if they are truly missing), gender, race, age at test date, matched pair grouping, test form, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth, birthweight, and an indicator if the pretest is from a previous year. * p<0.10, ** p<0.05, *** p<0.010

Online Appendix 1: Experiment Instructions

Time Preferences

Wave 1 (2010)

Now you are going to make some choices about candies. I will show you plates of candies and you will decide which plate you want. Some plates you choose, you can have TODAY, but some plates you choose you can have TOMORROW. I am going to put each plate you choose inside this box. At the end, you will CLOSE YOUR EYES and pick ONE plate from the box and that will be the plate you get to take home.

Okay, let's start! If you pick THIS plate (point to plate with 4), you could have it at the end of school TODAY. If you pick THIS plate (point to plate with 5), you could have it at the end of school TOMORROW.

Q1. Can you tell me, if you pick THIS plate (point to plate with 4), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it today.)
Q2. Can you tell me, if you pick THIS plate (point to plate with 5), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it tomorrow)

Okay, which plate do you want, this one TODAY or this one TOMORROW?

Okay, now I will put the plate you picked in the box. Let's play again! If you pick THIS plate (point to plate with 4), you could have it at the end of school TODAY. If you pick THIS plate (point to plate with 6), you could have it at the end of school TOMORROW.

Q3. Can you tell me, if you pick THIS plate (point to plate with 4), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it today.)
Q4. Can you tell me, if you pick THIS plate (point to plate with 6), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it tomorrow)

Okay, which plate do you want, this one TODAY or this one TOMORROW?

If you pick THIS plate (point to plate with 4), you could have it at the end of school TODAY. If you pick THIS plate (point to plate with 7), you could have it at the end of school TOMORROW.

Q5. Can you tell me, if you pick THIS plate (point to plate with 4), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it today.)
Q6. Can you tell me, if you pick THIS plate (point to plate with 7), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it tomorrow)

Okay, which plate do you want, this one TODAY or this one TOMORROW? Okay, now I will put the plate you picked in the box. Let's play again! If you pick THIS plate (point to plate with 4), you could have it at the end of school TODAY. If you pick THIS plate (point to plate with 8), you could have it at the end of school TOMORROW.

Q7. Can you tell me, if you pick THIS plate (point to plate with 4), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it today.)

Q8. Can you tell me, if you pick THIS plate (point to plate with 8), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it tomorrow) ' Okay,

which plate do you want, this one TODAY or this one TOMORROW? Okay, now I will put the plate you picked in the box.

Wave 2 (2012)

Now you are going to make some choices about candies. I will show you plates of candies and you will decide which plate you want. Some plates you choose, you can have TODAY, but some plates you choose you can have TOMORROW. I am going to put each plate you choose inside this box. At the end, you will CLOSE YOUR EYES and pick ONE plate from the box and that will be the plate you get to keep. Okay, let's start!
[TRIAL 1] If you pick THIS plate (point to plate with 3), you could have it at the end of our activity TODAY. If you pick THIS plate (point to plate with 3), you could have it TOMORROW.

Q1. Can you tell me, if you pick THIS plate (point to the TODAY plate), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it today.) Q2. Can you tell me, if you pick THIS plate (point to the tomorrow plate), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it tomorrow)

Okay, which plate do you want, this one TODAY or this one TOMORROW?

Okay, now I will put the plate you picked in the box. Let's play again!

[TRIAL 2] If you pick THIS plate (point to plate with 2), you could have it at the end of our activity TODAY. If you pick THIS plate (point to plate with 3), you could have it TOMORROW.

Q1. Can you tell me, if you pick THIS plate (point to plate with 2), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it today.) Q2. Can you tell me, if you pick THIS plate (point to plate with 3), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it tomorrow)

Okay, which plate do you want, this one TODAY or this one TOMORROW?

[TRIAL 3] If you pick THIS plate (point to plate with 1), you could have it at the end of our activity TODAY. If you pick THIS plate (point to plate with 3), you could have it TOMORROW.
Q1. Can you tell me, if you pick THIS plate (point to plate with 1), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it today.) Q2. Can you tell me, if you pick THIS plate (point to plate with 3), when can you have it, today or tomorrow? (Yes/No, if you pick THIS plate you can have it tomorrow)

Okay, which plate do you want, this one TODAY or this one TOMORROW?

Okay, now I will put the plate you picked in the box.

At the end of our activity today, you will pick the box that you get to keep.

Wave 3 (2013)

Now we are going to play some games where you make choices about candies. There are going to be 4 (show four on fingers) games. For each game, I will show you plates of candies and you will decide which plate you want. Some plates you choose, you can have TODAY, but some plates you choose you can have TOMORROW. In each game, the number of candies on the plates are different, so pay attention!

You will play 4 games (show four on fingers) but only 1 game (show one on fingers) will be the one that counts. At the end of all the games, I will roll a dice (show dice) to decide which game is the game that counts. If you chose the TODAY plate in the game that counts, you get that plate to take home today. If you chose the TOMORROW plate in the game that counts, you get that plate tomorrow. We don't know which game will be the one that counts until later, so when you make your decision in each game, you should pretend that it is the only game and that it is the one that counts!

[2 vs 3] (Put out 2 plates. Purple plate has 2 candies, Light blue plate has 3 candies.) Here are two plates. If you pick THIS plate (point to purple) you could have all the candies in here at the end of today. If you pick THIS plate (point to light blue) you could have all the candies in here at the end of TOMORROW.

Q1. Now before you pick - Can you tell me – if you pick THIS plate (point to red), when can you have it, today or tomorrow? (Yes/No) If you pick this plate you can have it TODAY
Q2. How about this one, when can you have it, today or tomorrow? (Yes/No) If you pick this plate you can have it TOMORROW.

Ok now we can decide. Which plate do you want, THIS plate for TODAY, or THIS plate for TOMORROW? Ok I'll write that down and we'll play the next game. (put away, mark response).

[2 vs 4] Put out 2 plates. Purple plate has 2 candies, Light blue plate has 4 candies. If you pick THIS plate (point to purple) you could have all the candies in here at the end of today. If you pick THIS plate (point to blue) you could have all the candies in here at the end of TOMORROW.

Ok now we can decide. Which plate do you want, THIS plate for TODAY, or THIS plate for TOMORROW? Mark response

[2 vs 5] Repeat as in [2 vs 4] Mark response. [2 vs 6] Repeat as in [2 vs 5] Mark response

Ok, great job. Now we're going to put those plates away, and at the end of all the games I will roll a dice to see which game will be the one that counts.

'Marshmallow' Task

Wave 1a (2010)

Now I'm going to go get your bags ready, I will be back in a little while and then we will get the pencils and stars out of the jar and pick the plate from the box, okay?

While you wait, I have some marshmallows here. I'm going to leave THIS one (one marshmallow plate) on the table here. If you want to, you can eat it now. BUT if you wait until I get back, then you can also have THIS one (two marshmallow plate).

Q1. Can you tell me how many marshmallows you get if you eat some BEFORE I come back? (Yes/No, you can have this one marshmallow)
Q2. Can you tell me how many marshmallows you get if you don't eat any UNTIL I come back? (Yes/No, you get ONE more marshmallow so you will have two)

Okay, I will be back in a while so we can play the pencil and candy game! Stay in your chair and I will be back.

Wave 1b (2010)

I have some marshmallows here. Which plate do you like better? (show plate with 2 and 4 mini-marshmallows). I'm going to leave to run an errand for 8 minutes. I'm going to leave this plate (show plate with 1 marshmallow) for you. Here are the rules. You can eat the marshmallow in this plate anytime you want to. If you wait for 8 minutes until I get back, you will get to eat this marshmallow (show plate with 1) AND the plate with these marshmallows (show plate with 4). But, if you want to eat this marshmallow before 8 minutes, you will get to eat this marshmallow (show plate with 1) AND the plate with these marshmallows (show plate with 2). Also, as soon as you eat the marshmallow, I will know and I'll come right back.

Q1. Can you tell me which plate I'm going to leave here? (Yes/No, I'm leaving this plate here – point to plate with 1 marshmallow
Q2. Can you tell me which plate you get if you eat this marshmallow before I come back? (Yes/No, you get this plate – point to plate with 2 marshmallow)
Q3. Can you tell me which plate you get if you don't eat this marshmallow until I come back in 8 minutes? (Yes/No, you get this plate – point to plate with 4 marshmallows)

Okay, I am going to come back in 8 minutes. You have to stay in your chair during the game or you won't get any prizes at the end. I can tell if you get off the chair, ok?

Wave 2 (2012)

I have some marshmallows here. (Show plates, each with 1 large marshmallow on them.) I'm going to leave to run an errand for 5 minutes. I'm going to leave this plate (show plate 1) for you. Here are the rules. You can eat the marshmallow in this plate anytime you want to. If you wait for 5 minutes until I get back, you will get to eat this marshmallow (show plate 1) AND the plate with this marshmallow (show plate 2). But, if you want to eat this marshmallow before 5 minutes, you will get to eat this marshmallow (show plate 1) and NOT the other one. Also, as soon as you eat the marshmallow, I will know and I'll come right back.

Q1. Can you tell me which plate I'm going to leave here? (Yes/No, I'm leaving this plate here – point to plate 1
Q2. Can you tell me which plate you get if you eat this marshmallow before I come back? (Yes/No, you get this plate – point to plate 1)
Q3. Can you tell me which plates you get if you don't eat this marshmallow until I come back in 5 minutes? (point to both plates).

Okay, I am going to come back in 5 minutes. You have to stay in your chair during the game or you won't get any prizes at the end. I can tell if you get off the chair, ok?

Wave 3 (2013)

Now we are going to play with these candies (show candies, unwrap them). These candies are for TODAY. I'm going to leave to run an errand for 15 minutes. I'm going to leave this plate (show purple plate with 1 candy) for you. Here are the rules. You can eat the candy in this plate anytime you want to. If you wait and don't eat the candy until I get back, you will get to eat this candy (show purple plate with 1) AND the plate with this other candy (show purple plate with another 1). But if you eat this candy before I get back, you will not get any more candies. Also, any time you want to end the game, raise your hand high and I will come back (demonstrate high hand raising) Remember, if you raise your hand, you will not get any more candies!

Q1. Can you tell me how many candies you get if you eat this candy or raise your hand, before I come back? (Yes/No) If you eat it or raise your hand, you will only get this candy (point) and not the other one (point)
Q2. What about if you wait to eat the candy until I get back – and you don't raise your hand – then how many do you get? (Yes/No) If you wait to eat the candy and you don't raise your hand, you will get this candy AND this candy.

(Leave plate with one unwrapped candy and take other unwrapped candy out of the room.) Okay, I am going to come back in 15 minutes. You have to stay in your chair during the game or you won't get any more candy. I can tell if you get off the chair, ok?

(Fill out response sheet. Watch child for 15 min or until ate / raised hand.)

Risk Preferences

Wave 1 (2010)

N/A

Wave 2 (2012)

In this activity, you will play for prizes. If you win stars, you will get to change the stars for prizes that you can take home. Do you see these prizes? These are the prizes you will be able to pick from. The token in this box always has ONE STAR on the bottom. See it? (pick up the token, show star sticker on bottom). If you pick the token in this box, you will get to exchange this token for a prize. The tokens in these boxes sometimes have TWO STARS on the bottom and sometimes have NO stars. See them? (pick up token with 2 stars, pick up token with no stars). If you get the token with TWO STARS, you can have two prizes. If you get the token with NO stars, you won't receive a prize for that token. You can't see if a token has a star until you pick it up. You will decide whether to pick out the box with the token that ALWAYS has ONE star, or pick out the box that has some tokens with TWO STARS and some tokens with NO STARS. You can pick EITHER this box OR this box. If you pick THIS BOX (show sure option), you will get this token that always has a prize (demonstrate). But if you pick THIS BOX (show risky option) there will be one more step. You're going to close your eyes and pick out one token. If the token you pick out has TWO STARS, you will get to change them for TWO PRIZES. But if the token you pick has NO STARS, you do not get any prizes. You will only pick out ONE token. You WONT KNOW whether the token has stars on the bottom until after you pick it.

Now you are going to make some choices about the boxes. I am going to put each box you choose inside this big box here. At the end, you will CLOSE YOUR EYES and pick ONE box and that will be the box you get to keep. Then, if it was a box with just one token, you get that token. But if it was a box with many tokens, some that have TWO STARS and some that have NO STARS, you will have one more choice – the choice of what token you want. You only pick out one token.

[Round 1] – 5 tokens in big box. Now we're going to play. I'm going to put this token HERE (put sure token with 1 star in the sure box, face down). I'm going to put these tokens HERE (put 2-star token plus 4 no star tokens in the risky box – make sure to demonstrate that only one has 2 stars and the others have nothing). Now I'm going to shuffle these around (shuffle the 5 token box around).

Q1. How many stars do you get if you pick THIS box (point to sure box). Answer: Yes (No), 1 star, and that equals 1 prize.
Q2. How many stars do you get if you pick THIS box (point to risky box). Answer: Yes (No), you sometimes get 2 stars, and sometimes NO STARS
Q3. What can you change a star for? Answer: Yes (No), each star equals 1 prize

Ok, I will put your box away and you will choose it out of the big box later. Let's go to the next one.

[Round 2] – 4 tokens in big box. I'm going to put this token HERE (put sure token with 1 star in the sure box, face down). I'm going to put these tokens HERE (put 2-star token plus 3 no star tokens in the risky box – make sure to demonstrate that only one has 2 stars and the others have nothing). Now I'm going to shuffle these around (shuffle the 4 token box around).

Q1. How many stars do you get if you pick THIS box (point to sure box). Answer: Yes (No), 1 star, and that equals 1 prize
Q2. How many stars do you get if you pick THIS box (point to risky box). Answer: Yes (No), you sometimes get 2 stars, and sometimes NO STARS
Q3. What can you change a star for? Answer: Yes (No), each star equals 1 prize. Ok, I will put your box away and you will choose it out of the big box later. Let's go to the next one.

[Round 3] – 3 tokens in big box. I'm going to put this token HERE (put sure token with 1 star in the sure box, face down). I'm going to put these tokens HERE (put 2-star token plus 2 no star tokens in the risky box – make sure to demonstrate that only one has 2 stars and the others have nothing). Now I'm going to shuffle these around (shuffle the 3 token box around).

Q1. How many stars do you get if you pick THIS box (point to sure box). Answer: Yes (No), 1 star, and that equals 1 prize
Q2. How many stars do you get if you pick THIS box (point to risky box). Answer: Yes (No), you sometimes get 2 stars, and sometimes NO STARS
Q3. What can you change a star for? Answer: Yes (No), each star equals 1 prize

Ok, I will put your box away and you will choose it out of the big box later. Let's go to the next one. [Round 4] – 2 tokens in big box. I'm going to put this token HERE (put sure token with 1 star in the sure box, face down). I'm going to put these tokens HERE (put 2-star token plus 1 no star tokens in the risky box – make sure to demonstrate that only one has

2 stars and the others have nothing). Now I'm going to shuffle these around (shuffle the 2 token box around).

Q1. How many stars do you get if you pick THIS box (point to sure box). Answer: Yes (No), 1 star, and that equals 1 prize.
Q2. How many stars do you get if you pick THIS box (point to risky box). Answer: Yes (No), you sometimes get 2 stars, and sometimes NO STARS
Q3. What can you change a star for? Answer: Yes (No), each star equals 1 prize

Ok, I will put your box away and you will choose it out of the big box later.

Wave 3 (2013)

Now we are going to play some games where you make choices about bouncy balls. First let me show you one of the spinner wheels we will play with. (Bring out spinner with 1/2 yellow). We will also play with these plates, which are for TODAY. (Bring out yellow and white plate; yellow plate has 2 balls white 0).

I'll spin and see what happens. (Spin wheel). Oh, it landed on (yellow/white).

(If it lands on yellow) If the wheel lands on YELLOW like this, that means you will get the balls on this plate (point to yellow plate 2 balls) OR (If it lands on white) If the wheel lands on WHITE like this, that means you will NOT get balls like this plate (point to white plate 0 balls)

Let's try again – this is just for practice. (Spin wheel). Oh it landed on (yellow/white).

Q1: What plate would you get? (Yes/No) You would get this plate – that's these/no balls (point)

Let's try again – this is just for practice. (show one more time).

Q2: What plate would you get?

Ok – are you ready to play for real now? There are going to be 4 (show four on fingers) games but only 1 game (show one on fingers) will be the one that counts. At the end of all the games, I will roll a dice (show dice) to decide which game is the game that counts. If you got balls in that game, you take those home. We don't know which game will be the one that counts until later, so when you make your decision in each game, you should pretend that it is the only game and that it is the one that counts! In each game you will decide whether to take a FOR SURE plate (bring out plate with 1, put on other side of spinner) or whether to SPIN. If you choose FOR SURE, then you get this plate, no matter what. We don't spin the wheel. But if you choose to SPIN, then we spin (point to wheel).

Let's start. [1/8] Put out Wheel with 1/8 yellow, FOR SURE plate and 2 / 0 plates.

Q3. If you decide to take the FOR SURE plate, which plate do you point to? (Yes/No) If you take FOR SURE, you get this plate TODAY

Q4. If you decide to SPIN, you will point to the wheel. Then what happens? (Yes/No) We spin the wheel, and you might get THIS plate or THIS plate TODAY.

Q5. When I spin the wheel, it could come up on YELLOW or WHITE (point) If it comes up

YELLOW, which plate do you get? (Yes/No) You get the plate with 2 balls (point to plate)

Q6. If it comes up WHITE, which plate do you get? (Yes/No) You get the plate with NO balls (point to plate)

Do you want the FOR SURE plate or do you want me to SPIN? (Accept verbal or pointing answers) - Mark response

[1/4] Put out Wheel with 1/8 yellow, FOR SURE plate and 2 / 0 plates. Ok, here is the next game. Now we are going to use this spinner. Do you want the FOR SURE plate or do you want me to SPIN? - Mark response

[1/2] Repeat -Mark response. [3/4] Repeat - Mark response

Social Preferences (Dictator Games)

Wave 1 (2010)

Now here is the next game! I am going to give you some candies, and you decide how many to keep and how many to give to Teddy. The candies are for taking home at the end of today. If you want to give candies to Teddy, you should put them on his plate right here. But if you want to keep candies, you should put them on your plate.

Q1. Can you tell me which plate is Teddy's? (That's right, this is Teddy's plate. Teddy will take the candies on his plate home.)
Q2. Can you tell me which plate is yours? (That's right, this is your plate. You will take the candies on your plate home.)

Okay, go ahead and put the candies on the plates you want. (set 5 candies in front of child).

Wave 2 (2012)

Now here is the next game! I am going to give you some stickers, and you decide how many to keep and how many to give to another student who didn't get to play the game today. The stickers are for taking home at the end of today. If you want to give stickers to the other student, you should put them on the other student's plate right here. But if you want to keep stickers, you should put them on your plate. The stickers you want to keep for yourself should go on this plate right here. You can give all, some, or none of your stickers to the other student, and you can keep none, some, or all of the stickers for yourself.

Q1. Can you tell me which plate is for sending? (That's right, this is the other student's plate. The other student will keep the stickers you put on this plate.)

Q2. Can you tell me which plate is yours? (That's right, this is your plate. You will take the stickers on your plate home.)

Okay, go ahead and put the stickers on the plates you want. (set 6 stickers in front of child).

Wave 3 (2013)

Here are some stickers, I am going to put these right here. They are yours now. (Put 6 stickers in a row in front of child, and point to child when saying 'you'. Also take out one

red and one blue plate.)

You are going to decide how many to keep, and how many to send to a boy/girl (match gender with gender of child) who did not get to play this game today and did not get any stickers. You don't know who the other boy/girl is and (s)he won't know who you are, but (s)he will get the stickers you send him/her in the bag here. You can keep as many stickers as you want, and you can send none, some or all of your stickers to the other boy/girl. Any stickers you are going to keep, you will put them on your plate, here. Any stickers you want to send to the other boy/girl, you will put them on the other child's plate, here. At the end of TODAY, you get to take home the stickers on your plate and the other boy/girl will get to have the stickers on his/her plate.

Q1. Can you show me which plate is your plate? This is your plate. You take home the stickers on this plate (point to blue plate) – we'll put them in your bag for you (show today bag).
Q2. Can you show me which plate is the other boy's/girl's plate? The other boy/girl takes home the stickers on this plate (point to red plate) – we'll put them in his/her bag (show other child's bag).

Okay, go ahead and decide where to put the stickers, on your plate or on the other boy's/girl's plate. Remember, it is up to you what you decide! (Mark response).

Great job, now I am going to put the stickers from your plate in YOUR BAG and the stickers from the other child's plate in HIS or HER bag.

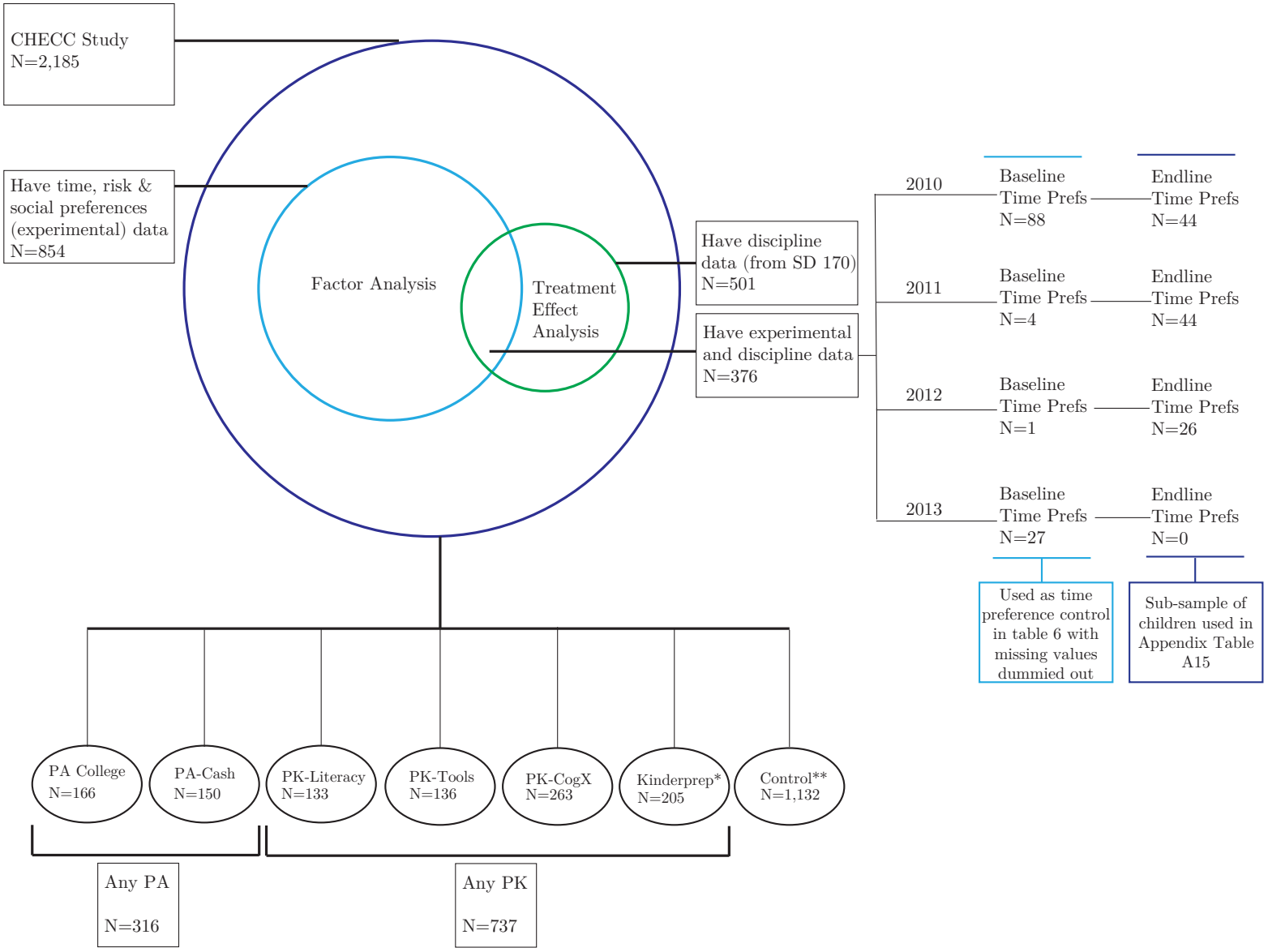
Online Appendix 2: Sample Development

Figure A1 provides a description of our overall sample and the analysis sample. CHECC enrolled 2,185 children ages 3-5 in four cohorts (2010, 2011, 2012 and 2013). Since the experiments were voluntary, only 854 of the children have experimental data. These children are used in our factor analysis. Five hundred and one children have disciplinary data because they are part of the school district (Chicago Heights) that agreed to share data with us. These are the children in the analysis capturing the association between early life skills and discipline. Since the experiments were conducted at various points in time of the intervention, separating our data into baseline and endline time preferences (as is done in the far right panel of Figure A1) shows that we have generally small numbers of children for whom we can evaluate the impact of the programs on time preferences.

The treatments that were part of the CHECC program are described in the bottom panel of Figure A1. Below, we outline exclusion rules for our analysis. First, we only include the children for whom we have disciplinary data, and do not use this overall sample. Second, our analysis of treatment effects uses the Parent Academy (PA), Preschool (PK) and control group data. We do not report on the treatment effects of the Kinderprep program because this is a shorter version of the Preschool program and we also do not find significant impacts of this program (this omits 73 children from the treatment effect analysis). Third, the 2011 cohort recruited an overly large group, and therefore at that time a decision was made to only actively follow up with a sub-set of these families. We only include this sub-set of control children in our analysis (this omits 70 additional children from the treatment effect analysis). Fourth, randomization was done by first creating matched groups of children based on age, gender and race, and then randomly assigning children from each group to treatment or control. Children for whom a matched group was not created were placed in the control group. As such, the regressions of treatment effects control for matched group and we omit children who do not have a matched group assignment (this omits an additional 129 children).

Our treatment effect analysis controls for baseline cognitive scores, baseline executive function scores and baseline time preferences (which are replaced with 0 if missing and include a dummy variable indicating truly missing values). We further control for the following factors in the treatment analysis: school attending, school year, grade level, year of randomization into the program, matched group (described above), whether the child was randomized into the program for two consecutive years, whether the child was at least 4 years old on September 1 of the year he/she enrolled in our program (which determines what year he/she left the program), test form (some tests have an A and B version), test maximum questions (some sub-tests asked more questions as the children got older), and the year from which the baseline assessment is taken (sometimes we have baseline scores up to a year prior to the randomization). When demographic controls are used, we control for gender, race, home language, age at test date, mother's age at child birth and child's birth weight. There are roughly 6,900 children from grades K to 8 in the Chicago Heights school district. One potential source of bias would be if the children for whom we have experimental and skill data are different in terms of misbehavior from the other children enrolled in the school district. Children for whom we do not have disciplinary referral data either never lived within the Chicago Heights School District (CHECC enrolled students from a number of districts in the area) or lived within the district but moved out prior to starting Kindergarten.

Figure A1: Sample Development

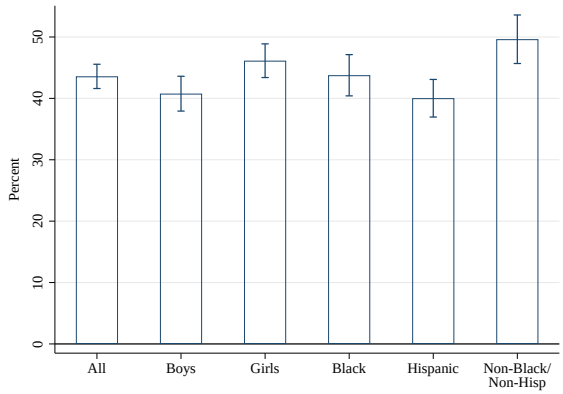


Note: * Kinderprep is excluded from the treatment effect analysis. **In 2011, a randomly selected sub-set of the control group is used in the treatment effect analysis

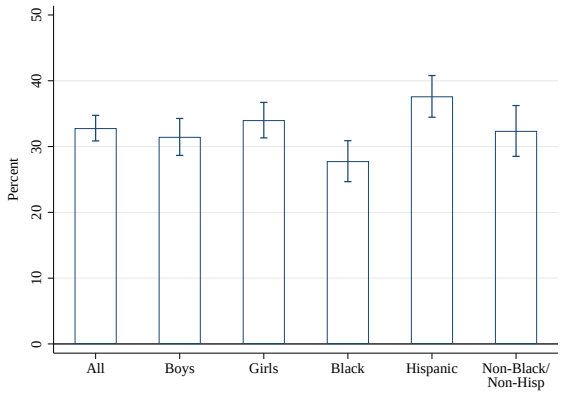
Online Appendix 3: Additional Analysis

Figure A2: Economic Preference Experiments

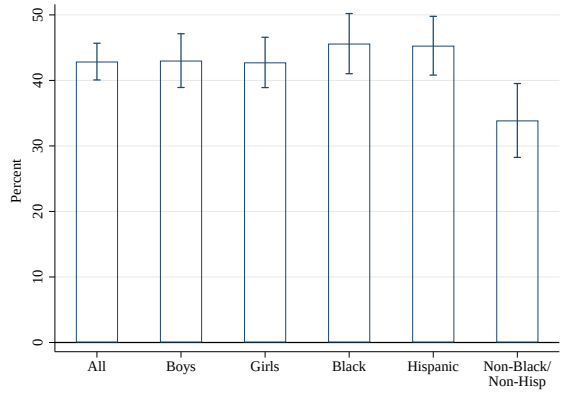
Risk Preferences
Percent of Safe Decisions
(a)



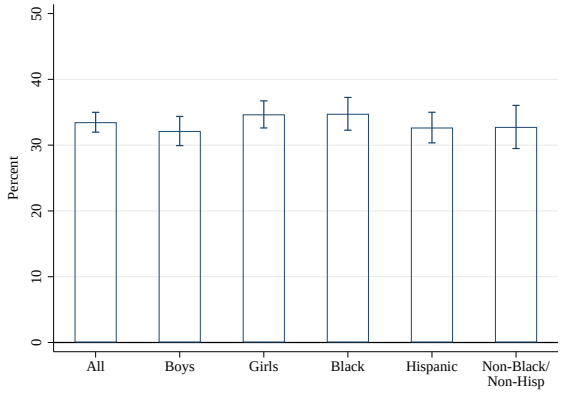
Time Preferences
Percent of Patient Decisions
(b)



Marshmallow Task
Percent who Ate Marshmallow Before 5 Minutes
(c)



Social Preferences
Percent who Shared
(d)



Note: This figure shows decisions in the economic experiments by gender and race of the child. (a) is the percent safe decisions in the risk experiment, (b) is the percent of patient decision in the time experiment, (c) is the percent of children who ate the marshmallow within 5 minutes in the marshmallow experiment and (d) is the percent who shared in the dictator experiment.

Table A1: Balance Tests of Differences in Disciplinary Referrals Comparing Children With and Without Experimental Data. Reporting Test Statistic for Each Academic Year of Data.

Panel A: Number of Disciplinary Referrals

Academic Year	t-test	p-value
2011	1.00	0.32
2012	-0.15	0.88
2013	-0.25	0.80
2014	0.58	0.57
2015	-1.43	0.15
2016	-0.29	0.78
2017	-0.61	0.54

Panel B: Received a Disciplinary Referral

Academic Year	t-test	p-value
2011	1.00	0.32
2012	0.23	0.82
2013	-0.45	0.65
2014	0.84	0.40
2015	1.02	0.31
2016	0.90	0.37
2017	0.74	0.46

Panel C: Severity of Disciplinary Referral

Academic Year	t-test	p-value
2011	-1.11	0.27
2012	-0.31	0.76
2013	0.65	0.52
2014	-0.66	0.51
2015	-0.79	0.43
2016	-0.98	0.33
2017	-1.10	0.27

Notes: This table reports p-values of t-tests comparing children with and without experimental data on disciplinary referrals by school year. Panel A reports p-values with respect to number of disciplinary referrals. Panel B reports p-values with respect to a binary variable indicating if a child has received at least one disciplinary referral in the school year and Panel C reports p-values with respect to the severity of the disciplinary referral as rated by independent coders.

Table A2: Description of Full Sample and Subsample with Disciplinary Referrals

Variable	Full Sample Mean/SD	Sub-Sample Mean/SD
Male (%)	47.85	51.60
Age of Child at Experiment (Months)	45.79	45.66
	(7.72)	(7.71)
Black (%)	37.05	21.54
Hispanic (%)	39.95	59.57
Birthweight (Lbs)	7.14	7.21
	(1.26)	(1.30)
Mother's Age at First Pregnancy (Years)	23.02	21.35
	(5.80)	(4.90)
Mother's Education		
Did Not Complete High School (%)	23.57	39.41
High School (%)	23.70	35.00
Some College (%)	42.59	34.12
Bachelor's (%)	13.18	6.18
Master's Degree and Above (%)	8.49	2.94
Household Income (Yearly, $)	28,875.32	20,909.63
	(22,037.10)	(14,835.73)
Number of Observations	854	376

Notes: This table reports the summary statistics for the full sample, i.e the sample for which we have the experimental, cognitive and executive function data and the sub-sample, i.e the sample that attended Chicago Heights School District, and therefore for whom we have disciplinary data. Standard deviation reported in parenthesis for continuous variables.

Table A3: Cognitive Skills, Executive Function and Economic Preferences - Subsample with Disciplinary Referrals

Cognitive Skills	Percentile Score	SD
Letter and Word Recognition (WJ-III)	32.65	26.82
Writing Skills (WJ-III)	31.53	27.19
Math - Applied Problems (WJ-III)	35.25	23.56
Math - Quantitative Concepts (WJ-III)	42.85	22.79
Receptive Vocabulary (PPVT-III)	26.84	23.74
Executive Functions	**Percent Correct**	**SD**
Attention (PSRA Assessment)	71.60	26.14
Emotion Regulation (PSRA Assessment)	57.67	24.35
Working Memory (Blair & Willoughby)	31.24	30.68
Inhibitory Control (Blair & Willoughby)	64.51	27.06
Economic Preferences	**Decisions**	**SD**
Time Preferences (% Patient Decisions)	35.35	34.20
Risk Preferences (% Safe Decisions)	43.36	33.73
Social Preferences (% Shared)	31.43	27.19
Delay of Gratification (% Ate Marshmallow)	44.51	48.46

Note: Cognitive measures are reported as percentile of age distribution. Executive function measures are reported as percent correct relative to total possible score. This table reports the data for the sub-sample for whom we have disciplinary data (N=376).

Table A4: Factor Loadings after Oblique Rotation

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Cognitive Skills				
Letter/Word Recognition (WJ-III Letter-Word Test)	**0.76**	-0.08	-0.13	0.13
Writing Skills (WJ-III Spelling Test)	**0.59**	0.20	-0.03	0.17
Woodcock-Johnston (Applied Prob.)	**0.68**	0.27	-0.04	-0.10
Woodcock-Johnston (Quantitative)	**0.85**	-0.13	0.07	-0.09
Peabody Picture Vocabulary Test (Receptive Vocabulary)	**0.59**	0.19	0.09	-0.22
Executive Functions				
Preschool Self-Regulation Assessment (Attention)	0.08	**0.78**	0.08	0.11
Preschool Self-Regulation Assessment (Emotion)	0.17	**0.55**	-0.02	0.16
Blair & Willoughby Assessment (Working Memory)	-0.08	**0.75**	-0.06	0.03
Blair & Willoughby Assessment (Inhibitory Control)	0.00	**0.77**	-0.05	-0.09
Economic Preferences				
Percent of Patient Decisions	-0.08	0.01	**0.80**	-0.15
Percent of Safe Decisions	-0.09	0.09	-0.04	**0.79**
Percent Shared	0.12	-0.05	**0.61**	0.40
Mischel Marshmallow Task (Ate Within 5 Min)	0.14	-0.40	-0.12	0.30

Notes: Factor loadings based on the exploratory factor analysis with direct quartimin rotation (Jennrich and Sampson, 1966) are shown. Factor loadings relating factors to corresponding potential dedicated measures are in bold. Based on 740 observations.

Table A5: Likelihood, Severity and Number of Disciplinary Referrals

	(1) 0-1	(2) 0-1	(3) Severity	(4) Severity	(5) Number	(6) Number
Percent of Patient Decisions [0,1]	-0.07*	-0.07**	-0.61*	-0.62*	-0.90***	-0.85***
	(0.04)	(0.03)	(0.36)	(0.33)	(0.34)	(0.31)
Percent of Safe Decisions [0,1]	-0.04	-0.04	-0.36	-0.42	-0.15	-0.29
	(0.04)	(0.03)	(0.36)	(0.32)	(0.38)	(0.33)
Percent Shared [0,1]	0.01	0.00	0.10	0.06	-0.05	-0.17
	(0.04)	(0.04)	(0.42)	(0.37)	(0.39)	(0.35)
Ate Marshmallow Before 5 Mins	0.01	0.00	0.15	0.02	0.31	0.17
	(0.02)	(0.02)	(0.22)	(0.22)	(0.22)	(0.21)
Executive Function Index	-0.02*	-0.02**	-0.20*	-0.21**	-0.19*	-0.23***
	(0.01)	(0.01)	(0.10)	(0.09)	(0.10)	(0.09)
Cognitive Index	-0.01	-0.01	-0.10	-0.09	-0.07	-0.12
	(0.01)	(0.01)	(0.15)	(0.13)	(0.15)	(0.14)
Male (d)	0.11***	0.11***	1.05***	1.00***	1.12***	1.10***
	(0.02)	(0.02)	(0.22)	(0.20)	(0.23)	(0.20)
Age at Experiment	0.00	0.00	0.03**	0.03**	0.03**	0.03**
	(0.00)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)
Black (d)	0.09*	0.12**	1.01***	1.14***	1.11***	1.30***
	(0.05)	(0.05)	(0.35)	(0.31)	(0.33)	(0.30)
Hispanic (d)	-0.01	0.01	0.17	0.31	0.11	0.29
	(0.03)	(0.03)	(0.31)	(0.26)	(0.30)	(0.27)
Observations	1265	1554	1313	1616	1317	1620
R2	0.14	0.13	0.10	0.09	0.12	0.12
Number of Children	313	369	318	375	318	375

Note: Individual-level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment included, but not shown. Columns (2) and (4) replace missing experimental data with mean values and add dummies for each missing experimental measure. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010

Table A6: Likelihood of Disciplinary Referrals - Disaggregated Economic Preference Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Available data					Completed data (accounting for item non-response)				
Percent of patient decisions [0,1]	-0.08**					-0.08**				
	(0.03)					(0.03)				
Percent of safe decisions [0,1]		-0.03					-0.03			
		(0.04)					(0.04)			
Percent shared [0,1]			-0.00					-0.00		
			(0.04)					(0.04)		
Ate marshmallow before 5min				0.00					-0.00	
				(0.02)					(0.02)	
Average of experimental measures					-0.11*					-0.10*
					(0.06)					(0.06)
Executive function index	-0.02**	-0.02**	-0.02*	-0.02	-0.02	-0.02**	-0.02**	-0.02**	-0.02**	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Cognitive index	-0.01	-0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Male (d)	0.10***	0.11***	0.11***	0.11***	0.11***	0.11***	0.11***	0.11***	0.11***	0.11***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Age at experiment	0.00	0.00*	0.00	0.00*	0.00	0.00	0.00*	0.00*	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Black (d)	0.09**	0.10**	0.11**	0.11**	0.09*	0.11**	0.13***	0.13***	0.13***	0.12***
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Hispanic (d)	-0.00	-0.01	0.00	-0.01	-0.01	0.01	0.01	0.01	0.01	0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Observations	1418	1425	1456	1384	1265	1554	1554	1554	1554	1554
R2	0.12	0.12	0.12	0.13	0.13	0.12	0.12	0.12	0.12	0.13
Number of children	342	342	351	338	313	369	369	369	369	369

Note: Individual-level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment included, but not shown. Columns (6)-(10) replace missing experimental data with mean values and add dummies for each missing experimental measure. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010.

Table A7: Severity of Disciplinary Referrals - Disaggregated Economic Preference Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Available data					Completed data (accounting for item non-response)				
Percent of patient decisions [0,1]	-0.73**					-0.71**				
	(0.33)					(0.32)				
Percent of safe decisions [0,1]		-0.34					-0.34			
		(0.34)					(0.34)			
Percent shared [0,1]			0.04					0.01		
			(0.37)					(0.38)		
Ate marshmallow before 5min				0.03					-0.04	
				(0.23)					(0.22)	
Average of experimental measures					-0.97*					-0.84
					(0.58)					(0.54)
Executive function index	-0.24***	-0.23**	-0.18*	-0.17*	-0.18*	-0.23***	-0.20**	-0.22**	-0.22**	-0.20**
	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Cognitive index	-0.07	-0.04	-0.10	-0.10	-0.09	-0.08	-0.08	-0.06	-0.06	-0.07
	(0.13)	(0.12)	(0.13)	(0.13)	(0.14)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Male	0.97***	1.00***	1.07***	1.04***	1.08***	1.00***	0.98***	0.99***	0.99***	1.03***
	(0.21)	(0.21)	(0.20)	(0.21)	(0.22)	(0.20)	(0.19)	(0.20)	(0.20)	(0.20)
Age at experiment	0.03**	0.03**	0.03**	0.03**	0.04**	0.03**	0.03**	0.03**	0.03**	0.03**
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Black	1.04***	1.08***	1.10***	1.09***	1.03***	1.10***	1.16***	1.17***	1.17***	1.17***
	(0.34)	(0.32)	(0.31)	(0.32)	(0.35)	(0.31)	(0.31)	(0.31)	(0.30)	(0.30)
Hispanic	0.18	0.13	0.23	0.17	0.16	0.27	0.27	0.29	0.29	0.34
	(0.29)	(0.28)	(0.28)	(0.29)	(0.31)	(0.27)	(0.27)	(0.27)	(0.27)	(0.27)
Observations	1477	1484	1509	1434	1313	1616	1616	1616	1616	1616
R2	0.09	0.09	0.09	0.09	0.10	0.09	0.09	0.09	0.09	0.09
Number of children	348	348	356	343	318	375	375	375	375	375

Note: Individual-level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment included, but not shown. Columns (6)-(10) replace missing experimental data with mean values and add dummies for each missing experimental measure. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010.

Table A8: Number of Disciplinary Referrals - Disaggregated Economic Preference Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Available data					Completed data (accounting for item non-response)				
Percent of patient decisions [0,1]	-0.97***					-0.93***				
	(0.31)					(0.31)				
Percent of safe decisions [0,1]		-0.17					-0.19			
		(0.33)					(0.32)			
Percent shared [0,1]			-0.28					-0.32		
			(0.35)					(0.36)		
Ate marshmallow before 5min				0.14					0.10	
				(0.21)					(0.21)	
Average of experimental measures					-1.40**					-1.33**
					(0.60)					(0.55)
Executive function index	-0.21**	-0.22**	-0.19**	-0.21**	-0.16	-0.25***	-0.23**	-0.23**	-0.24***	-0.22**
	(0.09)	(0.10)	(0.10)	(0.10)	(0.11)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Cognitive index	-0.11	-0.04	-0.10	-0.06	-0.04	-0.13	-0.08	-0.08	-0.08	-0.10
	(0.14)	(0.15)	(0.14)	(0.15)	(0.15)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
Male	1.06***	1.12***	1.21***	1.17***	1.10***	1.13***	1.13***	1.14***	1.15***	1.10***
	(0.21)	(0.21)	(0.20)	(0.21)	(0.22)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
Age at experiment	0.03*	0.03**	0.03**	0.04**	0.03**	0.03**	0.03***	0.04***	0.04***	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Black	1.10***	1.13***	1.21***	1.26***	1.13***	1.24***	1.29***	1.28***	1.32***	1.35***
	(0.32)	(0.31)	(0.31)	(0.31)	(0.32)	(0.31)	(0.31)	(0.30)	(0.31)	(0.30)
Hispanic	0.11	0.08	0.15	0.15	0.05	0.24	0.23	0.22	0.26	0.31
	(0.29)	(0.30)	(0.28)	(0.29)	(0.30)	(0.27)	(0.29)	(0.28)	(0.28)	(0.27)
Observations	1481	1488	1513	1438	1317	1620	1620	1620	1620	1620
R2	0.11	0.11	0.11	0.11	0.12	0.11	0.11	0.11	0.11	0.12
Number of children	348	348	356	343	318	375	375	375	375	375

Note: Individual-level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment included, but not shown. Columns (6)-(10) replace missing experimental data with mean values and add dummies for each missing experimental measure. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010.

Table A9: Likelihood, Severity and Number of Disciplinary Referrals - Additional Controls

	(1) 0-1	(2) 0-1	(3) Severity	(4) Severity	(5) Number	(6) Number
Percent of patient decisions [0,1]	-0.07**	-0.07**	-0.61	-0.63*	-0.93***	-0.85***
	(0.03)	(0.03)	(0.37)	(0.34)	(0.33)	(0.31)
Percent of safe decisions [0,1]	-0.03	-0.04	-0.26	-0.42	-0.10	-0.23
	(0.03)	(0.03)	(0.33)	(0.30)	(0.36)	(0.32)
Percent shared [0,1]	0.02	0.02	0.22	0.25	0.18	0.06
	(0.04)	(0.04)	(0.40)	(0.36)	(0.41)	(0.35)
Ate marshmallow before 5min	0.01	-0.00	0.09	-0.03	0.31	0.16
	(0.02)	(0.02)	(0.24)	(0.23)	(0.22)	(0.21)
Executive function index	-0.02**	-0.02**	-0.23**	-0.24***	-0.24**	-0.27***
	(0.01)	(0.01)	(0.11)	(0.09)	(0.10)	(0.09)
Cognitive index	-0.00	-0.01	-0.04	-0.04	0.01	-0.05
	(0.01)	(0.01)	(0.15)	(0.13)	(0.14)	(0.13)
Male (d)	0.11***	0.11***	1.09***	1.04***	1.09***	1.09***
	(0.02)	(0.02)	(0.21)	(0.20)	(0.22)	(0.20)
Age at experiment	0.00	0.00*	0.04**	0.03**	0.04**	0.03**
	(0.00)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)
Black (d)	0.08*	0.10**	1.01***	1.05***	1.09***	1.14***
	(0.04)	(0.04)	(0.35)	(0.32)	(0.34)	(0.31)
Hispanic (d)	-0.00	0.01	0.15	0.28	0.11	0.23
	(0.03)	(0.03)	(0.30)	(0.27)	(0.31)	(0.28)
Observations	1265	1554	1313	1616	1317	1620
R2	0.17	0.16	0.12	0.11	0.14	0.14
Number of children	313	369	318	375	318	375

Note: Individual-level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment included, but not shown. Columns (2) and (4) replace missing experimental data with mean values and add dummies for each missing experimental measure. Additional controls are birth weight, age of mother at first birth, mother's education, household income and variable dummies accounting for item non-response. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010.

Table A10: Likelihood, Severity and Number of Disciplinary referrals

	(1) 0-1	(2) 0-1	(3) Severity	(4) Severity	(5) Number	(6) Number
Average of time and social preferences	-0.08	-0.08*	-0.67	-0.64	-1.14**	-1.13***
	(0.05)	(0.05)	(0.49)	(0.45)	(0.44)	(0.41)
Percent of safe decisions [0,1]	-0.03	-0.04	-0.35	-0.46	-0.09	-0.30
	(0.04)	(0.03)	(0.34)	(0.32)	(0.36)	(0.33)
Executive function index	-0.02*	-0.02*	-0.20**	-0.19**	-0.17*	-0.21**
	(0.01)	(0.01)	(0.10)	(0.09)	(0.10)	(0.09)
Cognitive index	-0.01	-0.01	-0.08	-0.08	-0.07	-0.11
	(0.01)	(0.01)	(0.14)	(0.12)	(0.15)	(0.14)
Male (d)	0.11***	0.10***	1.04***	0.99***	1.14***	1.09***
	(0.02)	(0.02)	(0.21)	(0.19)	(0.22)	(0.20)
Age at experiment	0.00	0.00	0.03*	0.03**	0.03*	0.03**
	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)
Black (d)	0.08*	0.12**	0.97***	1.13***	1.00***	1.26***
	(0.04)	(0.05)	(0.34)	(0.31)	(0.32)	(0.30)
Hispanic (d)	-0.02	0.01	0.09	0.28	-0.00	0.24
	(0.03)	(0.03)	(0.30)	(0.26)	(0.30)	(0.28)
Observations	1310	1554	1360	1616	1364	1620
R2	0.12	0.13	0.09	0.09	0.11	0.12
Number of children	322	369	327	375	327	375

Note: Individual-level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment included, but not shown. Columns (2) and (4) replace missing experimental data with mean values and add dummies for each missing experimental measure. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010

Table A11: Likelihood, Severity and Number of Disciplinary Referrals - Disaggregated Cognitive and Executive Functions Measures

	(1) 0-1	(2) 0-1	(3) Severity	(4) Severity	(5) Number	(6) Number
Percent of patient decisions [0,1]	-0.06*	-0.07**	-0.54	-0.63*	-0.74**	-0.75**
	(0.03)	(0.03)	(0.37)	(0.34)	(0.34)	(0.33)
Percent of safe decisions [0,1]	-0.03	-0.04	-0.34	-0.51	-0.05	-0.25
	(0.04)	(0.03)	(0.36)	(0.32)	(0.39)	(0.33)
Percent shared [0,1]	0.02	0.02	0.20	0.23	0.07	0.02
	(0.04)	(0.04)	(0.38)	(0.34)	(0.39)	(0.34)
Ate marshmallow before 5min	0.01	-0.00	0.14	0.01	0.37	0.18
	(0.02)	(0.02)	(0.26)	(0.23)	(0.23)	(0.21)
Letter/word recognition (WJ-III letter-word test)	0.00***	0.00*	0.01***	0.01**	0.01	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Spelling (WJ-III spelling test)	-0.00	-0.00	-0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Woodcock-Johnston (applied prob.)	-0.00	-0.00	-0.01	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Peabody Picture Vocabulary test	-0.00	-0.00	-0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Woodcock-Johnston (quantitative)	-0.00	-0.00	-0.01	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Preschool self-regulation assessment (attention)	-0.14***	-0.14***	-1.38***	-1.27***	-1.98***	-1.83***
	(0.05)	(0.05)	(0.50)	(0.44)	(0.52)	(0.45)
Preschool self-regulation assessment (emotion)	0.11*	0.09*	0.74	0.63	0.54	0.32
	(0.06)	(0.05)	(0.61)	(0.49)	(0.55)	(0.45)
Operation span	-0.03	-0.05	-0.11	-0.36	0.27	0.21
	(0.04)	(0.04)	(0.41)	(0.37)	(0.41)	(0.38)
Arrows task	0.04	0.03	0.20	0.18	0.73	0.49
	(0.04)	(0.04)	(0.49)	(0.41)	(0.52)	(0.39)
Male (d)	0.10***	0.10***	1.02***	0.99***	1.07***	1.06***
	(0.02)	(0.02)	(0.20)	(0.20)	(0.21)	(0.20)
Age at experiment	0.00	0.00	0.03	0.03*	0.02	0.02*
	(0.00)	(0.00)	(0.02)	(0.02)	(0.02)	(0.01)
Black (d)	0.04	0.07	0.68*	0.83***	0.85***	0.94***
	(0.04)	(0.04)	(0.36)	(0.32)	(0.32)	(0.30)
Hispanic (d)	-0.01	0.01	0.07	0.24	0.03	0.16
	(0.03)	(0.03)	(0.29)	(0.26)	(0.29)	(0.27)
Observations	1265	1554	1313	1616	1317	1620
R2	0.18	0.17	0.13	0.12	0.15	0.14
Number of children	313	369	318	375	318	375

Note: Individual-level clustered standard errors in parentheses. Controls for year, school, grade, and year of experiment included, but not shown. Columns (2) and (4) replace missing experimental data with mean values and add dummies for each missing experimental measure. (d) for discrete change of dummy variable from 0 to 1. * p<0.10, ** p<0.05, *** p<0.010

Table A12: Average Treatment Effects on Number of Disciplinary Referrals

	(1) Number	(2) Number	(3) Number	(4) Number	(5) Number
PK	-0.08	-0.05	-0.24	0.01	-0.31
	(0.23)	(0.24)	(0.24)	(0.23)	(0.47)
PA	-0.38*	-0.36	-0.56**	-0.50**	-0.22
	(0.23)	(0.23)	(0.22)	(0.22)	(0.51)
Cognitive Baseline	-0.14	-0.17	-0.32**	-0.10	0.02
	(0.13)	(0.14)	(0.14)	(0.13)	(0.28)
Executive Function Baseline	-0.19*	-0.20*	-0.04	-0.10	0.39*
	(0.11)	(0.11)	(0.11)	(0.12)	(0.23)
Baseline Time Preferences	-0.17	-0.17	-0.26*	-0.16	-0.21
	(0.15)	(0.16)	(0.15)	(0.15)	(0.37)
Cognitive End Year		0.04	0.37***		
		(0.13)	(0.14)		
Executive Function End Year			-0.63***		
			(0.11)		
Cognitive End Kindergarten				-0.18	
				(0.15)	
Executive Function End Kindergarten				-0.47***	
				(0.14)	
Endline Time Preferences					-0.79***
					(0.16)
Observations	2318	2318	2318	2318	575
R2	0.127	0.127	0.142	0.137	0.235
Number of children	501	501	501	501	124

Notes: This table reports treatment effects on disciplinary behavior, clustered by individual using number of disciplinary referrals in a school year as the outcome variable. Columns 4 uses the same specification as column 1 but adds cognitive and executive function scores 1 year after treatment. Column 1-4 control for cognitive baseline scores and executive function baseline scores, baseline time preferences (with missing values replaced as 0 and including a dummy indicating if they are truly missing) gender, race, age at test date, matched pair grouping, baseline test form, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth and birthweight and an indicator if the pretest is from a previous year. Columns 3 and 4 omit endline executive function subtest form dummies due to collinearity. Column 5 omits baseline executive function subtest forms controls, indicator for aged above 4 at randomization and previous year pretest indicator to avoid collinearity due to small number of observations. $p<0.10$, ** $p<0.05$, *** $p<0.010$

Table A13: Average Treatment Effects on Skills - Disaggregated Treatment Arms

	(1) Cognitive Score	(2) Non-Cognitive Score	(3) Time Preference
PA-Cash	0.24*	0.09	-0.49
	(0.14)	(0.19)	(0.38)
PA-College	0.03	-0.03	-0.17
	(0.15)	(0.20)	(0.36)
PK-Tools	0.26*	0.34*	-0.07
	(0.15)	(0.20)	(0.38)
PK-Lit	0.47***	0.14	-0.82*
	(0.16)	(0.21)	(0.44)
Cog-X	0.30*	0.13	0.78*
	(0.16)	(0.21)	(0.40)
R2	0.78	0.67	0.32
N	376	376	152

Notes: This table reports disaggregated treatment effects on skills. PA Cash is the treatment arm where parents were randomized to receive part of the incentives as cash whereas PA College is the treatment arm where they would receive part of the incentives as a deposit in the child's post-secondary education account. PK Tools is the treatment arm under Preschool treatment where children received the Tools of the Mind curriculum and PK Lit is where children received the Literacy Express Curriculum. Cog-X is the treatment arm under Preschool treatment focused on cognitive and non-cognitive skills with a parent involvement component. Column 1 uses cognitive score at the end of the treatment year as the outcome (calculated as the mean of the standardized values of each of the subtests). Column 2 uses Executive Function score at the end of the treatment year, and Column 3 uses time preferences measured after treatment (measured as proportion of patient decisions). Regressions control for cognitive baseline scores and Executive Function baseline scores, gender, race, home language, age at test date, matched pair grouping, test form, number of previous assessments, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth, birthweight, assessor and an indicator if the pretest is from a previous year. * p<0.10, ** p<0.05, *** p<0.010

Table A14: Average Treatment Effects on Propensity to Have Disciplinary Referrals and Severity of Referrals - Disaggregated Treatment Arms

	(1) 0-1	(2) 0-1	(3) 0-1	(4) Severity	(5) Severity	(6) Severity
PA-Cash	-0.29*	-0.25*	-0.31**	-0.48*	-0.40	-0.50*
	(0.15)	(0.15)	(0.15)	(0.26)	(0.27)	(0.27)
PA-College	-0.19	-0.18	-0.29*	-0.37	-0.37	-0.57*
	(0.17)	(0.16)	(0.17)	(0.31)	(0.30)	(0.31)
PK-Tools	-0.13	-0.09	-0.13	-0.19	-0.14	-0.22
	(0.18)	(0.18)	(0.19)	(0.33)	(0.33)	(0.34)
PK-Lit	-0.03	0.02	-0.11	-0.12	-0.05	-0.27
	(0.17)	(0.18)	(0.18)	(0.30)	(0.31)	(0.33)
Cog-X	-0.27	-0.23	-0.27	-0.38	-0.30	-0.36
	(0.21)	(0.21)	(0.22)	(0.40)	(0.40)	(0.42)
Cognitive Baseline	-0.05	-0.05	-0.09	-0.10	-0.11	-0.16
	(0.07)	(0.08)	(0.08)	(0.15)	(0.16)	(0.16)
Executive Function Baseline	-0.10	-0.10	-0.03	-0.15	-0.16	-0.05
	(0.06)	(0.07)	(0.07)	(0.12)	(0.12)	(0.12)
Baseline Time Preferences	-0.07	-0.08	-0.11	-0.06	-0.09	-0.16
	(0.09)	(0.09)	(0.09)	(0.17)	(0.17)	(0.17)
Cognitive End Year		-0.00	0.12		0.01	0.20
		(0.09)	(0.09)		(0.16)	(0.16)
Executive Function End Year			-0.29***			-0.47***
			(0.07)			(0.13)
Observations	2264	2264	2264	2313	2313	2313
R2	0.167	0.170	0.185	0.104	0.105	0.113
Number of children	486	486	486	501	501	501

Notes: This table reports disaggregated treatment effects on disciplinary referrals. PA Cash is the treatment arm where parents were randomized to receive part of the incentives as cash whereas PA College is the treatment arm where they would receive part of the incentives as a deposit in the child's post-secondary education account. PK Tools is the treatment arm under Preschool treatment where children received the Tools of the Mind curriculum and PK Lit is where children received the Literacy Express Curriculum. Cog-X is the treatment arm under Preschool treatment focused on cognitive and non-cognitive skills with a parent involvement component. Column 1, 2 and 3 use a binary variable as the outcome indicating whether the child has had at least one disciplinary referral. Columns 4, 5 and 6 use the severity (rated by independent coders) of the disciplinary referral as the outcome variable. All regressions control for cognitive baseline scores and Executive Function baseline scores, gender, race, age at test date, matched pair grouping, test form, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth, birthweight, assessor and an indicator if the pretest is from a previous year. p<0.10, ** p<0.05, *** p<0.010

Table A15: Average Treatment Effects on Propensity to Have Disciplinary Referrals and Severity of Referrals - Endline Time Preferences

	(1) 0-1	(2) 0-1	(3) 0-1	(4) Severity	(5) Severity	(6) Severity
PK	-0.53	0.00	-0.45	-0.83	0.31	-0.44
	(0.35)	(0.30)	(0.33)	(0.61)	(0.57)	(0.56)
PA	-0.22	-0.15	-0.18	-0.52	-0.20	-0.27
	(0.31)	(0.30)	(0.33)	(0.60)	(0.55)	(0.62)
Cognitive Baseline	-0.76***	-0.05	-0.71**	-1.46***	-0.05	-1.32**
	(0.28)	(0.20)	(0.31)	(0.56)	(0.44)	(0.64)
Executive Function Baseline	0.37**	0.11	0.26*	0.64*	0.28	0.44
	(0.17)	(0.13)	(0.15)	(0.36)	(0.27)	(0.31)
Cognitive End Year	0.87***		0.87***	1.60***		1.60***
	(0.21)		(0.22)	(0.39)		(0.41)
Executive Function End Year	-0.59***		-0.53***	-0.97***		-0.82***
	(0.15)		(0.16)	(0.28)		(0.30)
Baseline Time Preferences		0.03	0.00		0.26	0.30
		(0.22)	(0.20)		(0.45)	(0.40)
Endline Time Preferences		-0.47***	-0.43***		-0.77***	-0.67***
		(0.12)	(0.14)		(0.25)	(0.24)
Observations	519	519	519	575	575	575
R2	0.327	0.304	0.349	0.206	0.188	0.216
Number of children	114	114	114	124	124	124

Notes: This table reports treatment effects on disciplinary behavior, clustered by individual using sub-sample of children who have endline time preferences. Columns 1,2 and 3 use a binary variable as the outcome indicating whether the student has had any disciplinary referrals. Columns 4,5 and 6 use severity, as rated by independent coders of the disciplinary referral. All regressions control for cognitive baseline scores, executive function baseline scores, baseline time preferences (with missing values replaced as 0 and including a dummy indicating if they are truly missing), proportion of patient decisions post-treatment, gender, race, age at test date, matched pair grouping, test form, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth, birthweight, assessor and an indicator if the pretest is from a previous year. * p<0.10, ** p<0.05, *** p<0.010

Table A16: Average Treatment Effects on Number of Disciplinary Referrals - Disaggregated Treatment Arms

	(1) Number	(2) Number	(3) Number	(4) Number	(5) Number
PA-Cash	-0.66**	-0.63**	-0.80***	-0.76***	0.25
	(0.28)	(0.27)	(0.27)	(0.27)	(0.59)
PA-College	-0.07	-0.07	-0.30	-0.20	0.64
	(0.28)	(0.28)	(0.28)	(0.27)	(0.57)
PK-Tools	-0.02	0.01	-0.06	0.11	1.26**
	(0.32)	(0.32)	(0.33)	(0.33)	(0.57)
PK-Lit	0.06	0.08	-0.28	0.07	0.58
	(0.29)	(0.30)	(0.31)	(0.28)	(0.56)
Cog-X	-0.23	-0.21	-0.38	-0.10	-2.02***
	(0.43)	(0.45)	(0.45)	(0.40)	(0.66)
Cognitive Baseline	-0.13	-0.15	-0.30**	-0.10	0.04
	(0.13)	(0.14)	(0.14)	(0.13)	(0.20)
Executive Function Baseline	-0.19*	-0.20*	-0.04	-0.10	0.47**
	(0.11)	(0.12)	(0.11)	(0.12)	(0.22)
Baseline Time Preferences	-0.15	-0.15	-0.24	-0.14	0.54
	(0.15)	(0.15)	(0.15)	(0.15)	(0.34)
Cognitive End Year		0.03	0.37***		
		(0.13)	(0.14)		
Executive Function End Year			-0.64***		
			(0.11)		
Cognitive End Kindergarten				-0.16	
				(0.15)	
Executive Function End Kindergarten				-0.44***	
				(0.14)	
Endline Time Preferences					-0.65***
					(0.17)
Observations	2318	2318	2318	2318	646
R2	0.129	0.129	0.143	0.138	0.241
Number of children	501	501	501	501	141

Notes: This table reports disaggregated treatment effects on disciplinary behavior, clustered by individuals using number of disciplinary referrals in a school year as the outcome variable. Columns 4 uses the same specification as column 1 but adds cognitive and executive function scores 1 year after treatment. Column 1-4 control for baseline cognitive and executive function scores, baseline time preferences (with missing values replaced as 0 and including a dummy indicating if they are truly missing), gender, race, age at test date, matched pair grouping, baseline test form, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth and birthweight and an indicator if the pretest is from a previous year. Columns 3 and 4 omit endline executive function subtest form dummies due to collinearity. Column 5 omits baseline executive function subtest forms controls, indicator for aged above 4 at randomization and previous year pretest indicator to avoid collinearity. $p<0.10$, ** $p<0.05$, *** $p<0.010$

Table A17: Average Treatment Effects on Propensity to Have Disciplinary Referrals and Severity of Referrals - IPW

	(1) 0-1	(2) 0-1	(3) 0-1	(4) Severity	(5) Severity	(6) Severity
PK	-0.15	-0.09	-0.17	-0.23	-0.14	-0.29
	(0.12)	(0.13)	(0.13)	(0.22)	(0.23)	(0.23)
PA	-0.26**	-0.25*	-0.33***	-0.45**	-0.42*	-0.56**
	(0.13)	(0.13)	(0.13)	(0.23)	(0.23)	(0.23)
Cognitive Baseline	-0.05	-0.03	-0.07	-0.10	-0.08	-0.13
	(0.08)	(0.09)	(0.09)	(0.15)	(0.16)	(0.16)
Executive Function Baseline	-0.11*	-0.11*	-0.04	-0.17	-0.17	-0.06
	(0.06)	(0.07)	(0.07)	(0.12)	(0.12)	(0.12)
Baseline Time Preferences	-0.05	-0.07	-0.11	-0.02	-0.05	-0.13
	(0.09)	(0.09)	(0.09)	(0.17)	(0.17)	(0.17)
Cognitive End Year		-0.03	0.10		-0.04	0.15
		(0.09)	(0.09)		(0.15)	(0.15)
Executive Function End Year			-0.30***			-0.48***
			(0.08)			(0.13)
Observations	2264	2264	2264	2313	2313	2313
R2	0.177	0.179	0.197	0.110	0.112	0.120
Number of children	486	486	486	501	501	501

Notes: This table reports treatment effects on disciplinary behavior weighted by inverse probability weights. Weights are estimated from running a probit regression on a binary variable indicating whether disciplinary referral data is available for the child regressed on gender, race and age at assessment of the child. Column 1, 2 and 3 use a binary variable as the outcome indicating whether the child has had at least one disciplinary referral. Columns 4, 5 and 6 use the severity (rated by independent coders) of the disciplinary referral as the outcome variable with same controls as described for Columns 1, 2 and 3 respectively. All regressions control for cognitive baseline scores and executive function baseline scores, baseline time preferences (with missing values replaced as 0 and including a dummy indicating if they are truly missing), gender, race, age at test date, matched pair grouping, test form, year of randomization, an indicator for aged above or below 4 at randomization, mother age at child birth, birthweight, and an indicator if the pretest is from a previous year. * p<0.10, ** p<0.05, *** p<0.010