An Evaluation of the Alief Independent School District Jump Start Program: Using a Model to Recover Mechanisms from an RCT [*]

Flávio Cunha[†], Andrea Salvati[††] and Kenneth I. Wolpin[†††]
March, 2020

Abstract

Recent research shows that the substantial differences in school readiness observed at the beginning of Kindergarten across socio-economic groups are partly due to disparities in the quality of children's early environment. Theory, consistent with a wide range of data, suggests that early interventions that target malleable, fundamental skills during sensitive periods of development in early childhood could help close these gaps. Indeed, empirical evidence shows that small-scale parenting interventions implemented by high-quality staff can lead to an improvement in parental investments and a boost in child development. Evidence about the impact of large-scale parenting interventions is more mixed. This paper reports the results of the evaluation of a parenting intervention developed and implemented by the Alief Independent School District in Texas. The goal of the intervention is to encourage and train parents to teach their children foundational skills for Pre-K. The results of a randomized controlled trial based on three yearly cohorts show that the program impacted parental investments and child development as measured by two different tests of school readiness. We go beyond reporting program impacts by building and estimating a model of parental choice of input levels. Our model allows for a production function of knowledge that features individual-specific coefficients that capture the marginal productivity of parental inputs. We find that the mechanism we posit for the program's impact is validated by the model estimates.

JEL Codes H31,I24,I25,I28,I3,J38,O1,O15,O54

---

[*] We thank the Family and Community Engagement Department of the Alief Independent School Districts for its support of the study, especially Shanceler Terry and Laura Escamilla. We also thank the Alief Family Liaisons for their help in recruiting families for the study and Yanitza Velez, our project coordinator. This research was supported from a grant by the Laura and John Arnold Foundation. Cunha also acknowledges support from the National Institutes of Health, grant 1R01HD073221-01A1.

[†] Professor of Economics, Rice University.

[††] Graduate Student, Rice University

[†††] Professor Emeritus of Economics, University of Pennsylvania; Professor of Economics, Rice University until July, 2020.

1. Introduction

Recent research shows that differences across socio-economic groups in school readiness at the beginning of Kindergarten remain large. Reardon and Portilla (2016) find that the gaps in reading and math achievement between high- and low-income families among a cohort of children born in 2010 are approximately one standard deviation. Moreover, such income gaps in achievement remain remarkably stable as children progress through school grades (Bond and Lang, 2012). The difficulty in overcoming these deficits in academic preparation after children start formal schooling is an important reason why researchers and policymakers alike have turned their attention to programs that aim to foster early childhood development.

Empirical research suggests that income gaps in early academic achievement are partly generated by disparities in the home environment. Despite progress in the last two decades, disadvantaged children still experience substantial deficits in the quantity and quality of early learning opportunities, whether they are measured in the time devoted to learning (Kalil et al., 2012), the availability of learning materials (Bossok et al., 2016), or engagement in routine activities that provide quality interaction with parents and other adults (e.g, family dinners, Putnam, 2015).

Research has shown that small-scale parenting interventions implemented by high-quality research staff can lead to an improvement in parental investments and a boost in child development (Suskind et al., 2013; Kalil et al. 2015, Cunha, Gerdes, and Nihtianova, 2019). In contrast, evidence about the impact of large-scale parenting interventions is mixed (Furstenberg, 2011). Holland et al. (2017) conducted an analysis of the Nurse Family Partnership (NFP) program in Memphis, TN.[1] They found that, by the end of the program (when children were two years old), mothers served by the nurses had higher levels of home investments, better parenting attitudes, and superior mental health. As a result, they report positive impacts for many measures of child development at age 6 and as late as age 12. However, such successes seem to be an exception in this literature. For example, Love et al. (2005) found that the home visitation component of Early Head Start achieved modest impacts on family investments or child development. Similarly, St. Pierre et al. (2003) report that the Department of Education's Even Start Program (at a cost of $ $13,674 per family) yielded no impacts on parental investments or child development outcomes. Evaluations of the Home Instruction for Parents of Pre-school Youngsters (HIPPY), a widely adopted two-year home-visitation program (in over 100 locations in the U.S. and in 14 countries) costing between $1,500 to $2,000 per family per year, also have found mixed results. Although one small-scale randomized trial of 69 HIPPY

[1] The NFP is a home visitation program for eligible first-time mothers beginning during pregnancy and extending through the first two years of the child's life. The program began in the 1970's and has served several hundred thousand women.

participants produced substantial improvements in test scores, the results were not replicated in a randomized trial of similar families who entered the program the following year (Baker, Piotrkowski, & Brooks-Gunn, 1998, 1999).

Individual school districts have taken different approaches. For example, the Houston Independent School District (HISD) has adopted the HIPPY program. In contrast, the Alief Independent School District (AISD), a school district that neighbors the HISD, designed its own program, called Jump Start, that is provided at 10 percent of the cost of the HIPPY program. In contrast to the general goal of the HIPPY program, the JSP is aimed specifically at preparing three-year old children for entry into the AISD Pre-Kindergarten program. In addition, the JSP is center-based, providing parents with training and materials to work with their children in their home.

In this paper, we provide an evaluation of the Jump Start Program (JSP) conducted at the request of the district. The evaluation is based on a randomized controlled trial of three separate JSP cohorts, beginning with the 2016-2017 academic-year cohort. The purpose of the RCT was to provide the district with the best evidence about the impact of the JSP. However, we go beyond the original purpose of the evaluation to try to understand the mechanism generating the program's impact. To that end, we develop and estimate a model of parental behavior in providing inputs into the production of children's age-relevant cognitive skills. A novel feature of the model is that, unlike in the extensive literature on the estimation of child cognitive achievement production functions, we allow for the marginal productivity of inputs to be heterogeneous across parents.[2] The JSP is modeled as potentially shifting the distribution of parental productivity. Removing the assumption that all parents have the same level of productivity raises challenging issues in the estimation of the production function parameters. The reason is that, as demonstrated in the optimizing model, inputs chosen by parents will depend on their parental productivity.

The RCT evaluation of the JSP included families living in the catchment areas of all of AISD's 24 elementary schools.[3] Overall, during the three-year evaluation period the number of families included in the program averaged about 200 per year, with an additional 200 in the control group. Families were offered 25 dollars to participate in the study. Depending on the year, between 90-95 percent of those registering for the Jump Start Program agreed to be in the study. All families that registered for the program were included

---

[2] This setting falls into the general framework of endogenous regressors with heterogeneous effects considered in Florens et al. (2008).

[3] Prior to the evaluation, the JSP was limited to about 100 families in a subset of the schools. The expansion of the program, as well as the implementation of the evaluation study, was funded by a grant from the Laura and John Arnold Foundation.

in the lottery randomization regardless of whether they agreed to be in the study. Only those in the study are included in the evaluation.

All children enrolled in the study were given assessments at the time the families were recruited into the program and at the end of the program. One assessment, designed by AISD staff, is based explicitly on the JSP curriculum. The second, the Bracken School Readiness Assessment, is a nationally normed test of five concepts, only one of which coincides with the Jump Start curriculum content. The additional assessment, specific to the evaluation, provides evidence on whether, due to the Jump Start training, parents choose to teach their children more advanced concepts and skills than those in the Jump Start curriculum. In addition to the 25 dollars families were offered to participate in the study when they applied for the program, families were offered 50 dollars to participate in the post-program assessments eight month later. Over the three cohorts, 79 percent of the families that originally agreed to be in the study also agreed to remain in the study. Not only was attrition low, but the treatment and control families in the post-program sample did not differ in observable characteristics.

The importance of a comparison group in evaluating the JSP is especially clear in the case of early childhood interventions. Young children are continually learning new things. In the case of the material covered in the Jump Start curriculum, the test score of the control group children increased from 50 percent at the baseline to 68 percent at the end of the program eight months later. Thus, even if the JSP had no impact, one would expect an 18 percentage point increase in the score of those in the program. Based on a (modified) difference-in-difference regression, the gain in the score of the children in the program was 7 percentage points higher than for the control group children. Program gains were particularly large for two of the Jump Start modules, name recognition and book handling, with the treatment effect being over 20 percentage points for each.

The effect of the program on Bracken test scores were considerably smaller in absolute terms. Over the program period, the Bracken score increased from 23 to 34 percent for the control group. Based on a (modified) difference-in-difference regression, the JSP increased the Bracken score by a little over 2 percentage points. Using a different metric, the JSP reduced the percent of children defined by the Bracken score as either delayed or very delayed by 4 percentage points, about double the reduction for the control group (about one-half of the control group children were in this category at the baseline). Thus, there was some spillover in learning beyond the JSP curriculum.

In addition to the child assessments, parents enrolled in the evaluation study also completed a questionnaire both at the time of the recruitment and at the end of the program period. Besides the usual demographic information, measures of parental interactions were also obtained in an attempt to understand the mechanisms behind the treatment impact. Data were collected on the number of days the parent read to the child, and on whether family members helped the child learn to recognize letters, colors, numbers and shapes. At the baseline, on average, control-group parents read to the child three days a week and between 80 and 90 percent of the families reported helping their child recognize letters, colors and numbers and 67 percent reported helping with shapes. Control-group families increased their average number of days spent reading to the child by a little over one day, and the other activities increased to 89, 99, 96 and 83 percent. Because family engagement in these latter activities became almost universal, there was essentially no difference between the treatment and control groups at the end of the program period. However, the number of days spent reading to the child increased by about one-half a day more for the treatment group.

Although JSP parents read more frequently to their child than did control parents, the relevance of this parental activity as a mechanism underlying the JSP treatment effect depends on the extent to which the JSP induced increase had an effect on test scores. There is evidence for such an effect in the literature on educational production functions (see, for example. Todd and Wolpin (2007)). To ascertain the importance of this potential mechanism, we posit a model in which a parent decides on the frequency of reading to their child over some period of time, in this case from the start to the end of the JSP. The parent is assumed to care about the child's ending knowledge, which is determined by the child's initial knowledge and the parental input of reading time (and other inputs).

Adopting an explicit choice model allows for a potentially important extension of the formulation of achievement production functions in the context of the early child development and educational production function literatures.[4] Those literatures have universally assumed that all parents, in the case of parental inputs, are equally productive in producing child development outcomes. Instead, we assume that parents may be heterogeneous in the marginal product of the time spent reading to their child (the one input, as noted above, that varies in the sample and that was affected by the JSP). We show how the RCT supplies exogenous variation that enables us to identify the parameters of the model without making a distributional assumption about the heterogeneous marginal product parameter (or any other error terms). The effect of the JSP is modeled as changing the parent marginal productivity distribution, which, in turn, optimally affects the amount of time parents spend reading to their child. We estimate the model using empirical likelihood based on matching first- and second- moments of the data to their theoretical counterparts in the model, and restricting estimation to the first two moments of the marginal productivity parameter.

[4] Todd and Wolpin (2003) draw a distinction between the early childhood development (ECD) literature and the educational production function (EPD) literature. Although the former focuses on parental inputs prior to formal education and the latter on school inputs, both adopt a production function framework.

Our identification analysis demonstrates that the marginal product distribution (the mean and variance) is potentially only set identified, there are either one or two values of the mean and variance that satisfy the moment conditions. We find that two values are consistent with the data. In one case, the mean and variance of the marginal product parameter, for both the treatment and control groups, are statistically significant. In the other case, the means are statistically significant but not the variances, indicating that homogeneity is not rejected. For both sets of estimates, the JSP is found to increase the mean of the distribution, consistent with the mechanism we posit for the JSP.[5]

We also show how the standard approach of estimating the achievement production function can be interpreted only through the lens of a model. For example, under the interpretation of our model, the Wald estimator using the treatment assignment as an instrument for the input, not only doesn't identify the mean of the marginal product distribution, but actually is not even a function of the marginal product distribution parameters.

Our paper connects with the literature on the evaluation of parenting programs as well as with the estimation of mechanisms. Two papers are particularly relevant. First, we build on the econometric analysis of Florens et al. (2008).We estimate a production function of human capital formation that allows for individual heterogeneity in coefficients that capture the marginal productivity of inputs. Florens et al. (2008) present a richer and more general analysis of identification, but do not implement any estimation.

Second, our paper is related to recent work by Attanasio et al. (2019), who aim to uncover the short-term impacts of the Jamaica home visitation program as implemented in a larger scale in Colombia. The paper also takes advantage of data collected as part of a randomized controlled trial to estimate a production function of cognitive and socio-emotional skills that can differ between treatment groups. Their finding differs from ours in that they cannot reject that the treatment and control families are adopting the same production function. An important difference from our analysis is that we allow for individual-specific production function coefficients.

The rest of the paper is organized as follows. The next section briefly describes the Alief Independent School District Jump Start Program. Section 3 discusses the evaluation study, including the experimental design, the recruitment and the randomization. Section 4 presents the results from the evaluation study. Section 5 develops the details of the analysis to uncover the JSP's mechanism, including the model, identification, estimation, and results. The final section concludes.

[5] The variance of the marginal product parameter is also increased by the JSP for the estimate in which the variances are statistically significant.

2. The Jump Start Program

The JSP was created and designed by staff from AISD's Department of Family and Community Engagement (FACE). AISD is located in southwest Houston, Texas. In the 2017-18 academic year, AISD served over 46,000 students, 83% of which were economically disadvantaged and 43% of which were English Language Learners. To be eligible for the JSP, parents must live within the AISD catchment area and must have a child who is between 36 and 47 months old on September 1 of each academic year. The characteristics of the JSP families reflect the composition of the district. About 50 percent of the families that apply to participate in the JSP receive food stamps, 57 percent of the parents are Hispanic and their mean years of completed schooling is about 12.

The cost of the JSP, primarily for the materials, is less than $200 per family. The twenty-two week curriculum was developed by the district specifically to be aligned with the district's pre-K program to ensure that children have the necessary foundational skills upon entering pre-K. The week-by-week Jump Start curriculum, which includes learning colors, acquiring fine motor skills, counting, name recognition and book handling, is summarized in Appendix A. Three times each month, over an eighth-month period, parents meet for one hour with a trained family liaison, a paraprofessional that administers the program, at their local elementary school in a group setting (there is one additional meeting at the end of the program year). The children are present at the third meeting of each month to enable the family liaison to assess parent fidelity to the curriculum. Thus, Family Liaisons guide parents in how to teach their children the skills necessary for their success in the AISD Pre-K program.

3. The Evaluation Study

The evaluation study of the JSP began with the 2016/17 school-year cohort and continued for an additional two years, the 2017/18 and 2018/19 school-year cohorts. Prior to the start of the study, about three-quarters of the 24 district elementary schools provided the JSP to families within their catchment area. Total enrollment throughout the district amounted to about 100 families. To participate in the program, each school had to cover the cost of the program out of their own budgets, $150 for materials for each family and the time of the Family Liaisons. The cost of the program to the schools was a major impediment to a school adopting Jump Start or expanding it to more families. As part of the evaluation study, the Laura and John Arnold Foundation provided the money needed to expand the program to all 24 elementary schools in all three years.

3.1. Recruitment

Prior to the study, AISD family liaisons recruited families on a first-come first-serve basis. Recruitment ended when a pre-set target number of families, determined by the amount of money each individual school had allotted to the program, was reached. To implement the randomized controlled trial, the process of recruiting families for the study was changed.

In the first step, at the recruitment meeting with the family liaison at the school, the family representative was informed that there was an oversubscription to the program and that, to be fair, participants in the program would be chosen through a lottery. Any family desiring to be in the program had to agree to be in the lottery.[6] After agreeing to be in the lottery, a Jump Start application form was completed by the person who came with the child, almost always the child's mother, and the child was administered a baseline assessment, the JSP test designed by AISD, that covers the content of the JSP curriculum. The family was then informed that Rice University was engaging in a study of the JSP to assess whether the program was achieving its aims. The family was told about their role in the study and offered 25 dollars as an incentive to participate. Families that agreed to participate were consented pursuant to IRB regulations. Participation in the study required the completion of a survey instrument and the administration of another child assessment, the Bracken School Readiness Assessment, a nationally normed test of five concepts, only one of which coincides with the Jump Start curriculum content.

The recruitment strategy in the first year of the study was based on the assumption that interest in the Jump Start Program was sufficiently great in all of the 24 school catchment areas that a recruitment target of 20 families per school could be met. With the randomization, 10 families would be assigned to the treatment sample (the JSP participants) and 10 to the control sample (the program non-participants). Twenty of the twenty-four schools reached the 20-family enrollment target for the lottery; in all, 461 families enrolled in the lottery (see table 1). Although the first-year recruitment in terms of the enrollment numbers was clearly successful, in part because family liaisons were given a monetary incentive to reach the 20-family goal, there was not uniformly high interest among the families across the schools. In half of the schools, at least one family who received a spot in the Jump Start Program through the lottery never attended a Jump Start session. Overall, 10 percent of the families never attended a Jump Start session and 25 percent attended five or fewer sessions out of the 22 total sessions (see table 4).

In response to this outcome, the recruitment strategy was modified in the second and third years of the study. Liaisons were provided a monetary incentive if they recruited 14 families and a group incentive if they recruited additional families. In the second year, in total 389 families entered the lottery spanning the 24 schools. In the third year, two family liaisons left the district just before the start of the academic year

[6] We are not aware of any families declining the program because of the lottery.

and were not replaced. Over the 22 remaining schools in the third year, 366 families entered the lottery. In contrast to the first year, in the second year only 4 schools had a family that never attended a Jump Start session, accounting for only 2 percent of the families, and only 13 percent attended fewer than 5 sessions. However, in year three, for reasons that are unclear, attendance declined; 8 percent of the families never attended a session and 19 percent attended five or fewer sessions, not quite as problematic as the first year, but significantly worse than the second year.

The results of the recruitment are shown in table 1. Participation in the study was high in all three years. In the first year, 89.5 percent of control families and 90.9 percent of treatment families that agreed to be in the Jump Start lottery also consented to be in the evaluation study.[7] There was even greater participation the second and third years, 93.7 percent of the control families and 92.5 percent of the treatment families in year 2 and 96.0 percent of the control families and 94.2 percent of the treatment families in year 3. Cumulatively over the three years, 1,216 families entered the Jump Start lottery of which 1,126 participated in the study, a refusal rate of only 7.4 percent.

Table 1 also shows the results of the end-of-year (about 8 months after the start of the program) recruitment. As seen, attrition from the study fell in each year. Noting that we attempted to contact only those families that had agreed to be in the study at the baseline, 71.3 percent of them in the first year participated in the end-of-year re-interview, 82.3 percent in the second year and 84.8 percent in the third year. More importantly, in the last two years, the rate of attrition from the study, which is higher for the control group, fell significantly more for the control group than for the treatment group. In year one, the attrition rate (refusals and contact failure) was 18.5 percentage points higher for the control group. In years 2 and 3, the difference in the refusal rate fell to 9.4 and 6.0 percentage points.

3.2. Randomization

The randomization was conducted using a block design, the same in all three years. Families were divided into four groups based on whether their score on the Jump Start baseline test was above or below the mean of the lottery sample and on whether the child's age in months was above or below the mean of the lottery sample. Families were randomly selected within each school and blocking group. The number of families chosen within each school to participate in the program was constrained to be one-half of the total number of lottery families in the school, with the qualification that when the total number of families was an odd number the additional family was placed into the treatment group.

[7] Recall that all families, regardless of whether they agreed to be in the study, were entered into the lottery.

Table 2 provides information on the outcome of the randomization in each year. The table excludes families that did not consent to be in the study (consistent with IRB regulations). Because all of the results of the study are based on the sample of study participants, it is most relevant to document differences in characteristics between the treatment and control groups for that sample. For each variable shown in the first column of the table, the second column shows it's mean in the study sample (treatment plus control families), the third column the difference in the value of the variable between the treatment group and the control group and the last column the p-value associated with the hypothesis that the difference shown in column three is equal to zero. The first two variables, the child's baseline Jump Start test score and the child's age, are the blocking variables; the rest of the variables were not used in the randomization. As seen in the table, the differences between the treatment and control samples are small in magnitude (relative to the means) and have p-values well above conventional levels of statistical significance.

Table 3 shows the differences in the same characteristics of treatment and control groups as in table 2 for the post-program study sample. As in the baseline study sample, these differences are small and have p-values well above conventional levels of statistical significance. Comparing table 2 and 3, the three-year pooled sample of treatment and control families is almost identical in mean characteristics; the only slight difference is that Hispanics comprise about 3 percentage points more of the post-program sample than of the baseline sample.

3.3. Jump Start Program Attendance

As noted, Jump Start lessons are held in group sessions within each school three times a month with 22 lessons in total over the school year. The first two lessons each month are with the parents alone and the third with the parents and their children. Parents who miss either or both of the first two lessons in each month may make them up at another time. Parents who miss 3 lessons in a row, without make-up, are dropped from the JSP.

Table 4 shows selected statistics in each year from the distribution of completed (inclusive of made-up classes) weeks for the treatment group over the 22 weeks of the program. The mean number of weeks completed was 13.5 in year one, increased to 17.4 in year 2 and dropped slightly to 16.1 in year 3. These differences are also reflected in graduation rates, which require that a family complete 16 or more weeks out of the 22. In large part due to the recruitment strategy in year one, only 57.2 families completed the Jump Start program in that year. Given the change in the recruitment strategy, 74.4 percent completed the program in year 2 and 71.6 percent in year 3. In years 1 and 2 about half of the families that didn't graduate completed 5 weeks or less, while that is true for about two-thirds of the families in year 3. And, among the non-graduates, the mean number of weeks completed was 4.5, 6.8 and 4.4 for the three years.

4. Data

4.1. Baseline Test Score Summary Statistics

Table 5 provides summary statistics of the baseline Jump Start and Bracken raw test scores (percentage correct answers) for families in the post-test sample for the three years separately and pooled. The mean Jump Start test score pooled over the three years is 51.0 percent. The range over the three years is only 5.1 percentage points, from a low of 48.8 percent in year 3 to a high of 53.9 percent in year 2. The standard deviation of the score is about 26 percentage points. Approximately 20 percent of the sample has a test score below 25 percent and another 20 percent of the sample 75 percent or higher.

Bracken baseline test scores are considerably lower than Jump Start test scores. As seen, the mean baseline score on the Bracken is 23.9 percent averaged over the three years. As with the Jump Start test, the range over the three years is narrow, 22.9 to 25.1 percent. The standard deviation of the pooled sample is 19 percent. The Bracken is a nationally normed test. The children of families that applied to participate in the Jump Start program are well below the national mean. The mean percentile score, averaged over the three years, is only 27 and about 50 percent of the children are judged to be delayed or very delayed in their cognitive development.

4.2. The Impact of Jump Start on Test Scores

Tables 6 and 7 present estimates of the impact of the Jump Start program on the Jump Start and Bracken test scores using the three-year pooled sample. The first column in each table reports the end-of-year difference in the test scores of the treatment and control families, the second the difference-in-difference estimate and the third the difference-in-difference estimate conditioning on the initial baseline test score. The first two estimates assume that the difference in test scores between the treatment and control children that did not take the end-of-year assessment would have been the same as the difference for those children that did take the end-of-year assessment. The estimate in the last column of each table assumes that treatment-control difference of those children not assessed would have been the same as for those that were assessed, conditional on having the same baseline test score.

The first row of each table reports the effects for the overall test (all modules), while the rest of the rows report effects for the separate test modules. Beside each nodule in parentheses is the number of test questions and in the brackets, the baseline and end-of-year test scores for the control group. As seen, there are 26 questions in the Jump Start test. The test includes 6 modules, 10 questions on recognizing colors, 8 testing fine motor skills (for example, coloring a circle inside the lines), 2 on counting, 1 on sorting objects that are similar, 1 on recognizing one's printed name and 4 on book handling (for example, recognizing the title of a book).

The control-group mean baseline score for the overall test is 50.4 percent. The highest baseline score is obtained on the colors module, 58.4 percent, followed by the fine motor skills module, 56.8 percent. At the other extreme, only one-quarter of the children were able to recognize their printed name and most children were able to answer only one of four questions about book handling. The counting and sorting modules lie in between, a mean score of around 40 percent.

As should not be surprising, children in the control group learn without participating in the Jump Start program. Indeed, the overall score for the control group increases by 18.3 percentage points, from 50.4 to 68.7 percent. The increases for the individual modules are: colors - 16.3 percentage points, Fine Motor Skills – 18.8 percentage points, Counting – 22.4 percentage points, Sorting – 24.3 percentage points– Name Recognition – 23.3 percentage points, Book Handling – 17.3 percentage points. Clearly, any before-after comparison for the children participating in the Jump Start program would need to account for this growth.

As seen in table 6, the three estimates of the impact of the Jump Start program are similar in magnitude. Concentrating then on the last column, the Jump Start program is estimated to increase the overall test score by 7.0 percentage points more than the increase in the test score of the control group (p-value of .001). Among the modules, the impact of the program is estimated to be 3.2 percentage points (p-value= .04) for Colors, 4.0 (p-value=.004) for Fine Motor Skills, 2.8 (p-value=.30) for Counting and 6.4 (p-value=.04) for Sorting. Most striking is the impact on the score for name recognition, 22.2 percentage points (p-value=.001), and the impact on the score for book handling, 21.5 percentage points (p-value=.001). Thus, for example, although less than one-half of the control group children can recognize their name on the post-test, about 70 percent of the treatment group can do so.

Table 7 shows the results for the Bracken test. The Bracken test consists of 85 questions contained in five modules that test knowledge of colors, letters, numbers, sizes and shapes; testing in each module ends when the child has answered three consecutive questions incorrectly. Recall that the overall score on the baseline tests was only 23 percent, less than half of the score on the Jump Start test. The percentage point gain in the Bracken test score for the control group was also smaller than that of the Jump-Start test, 11.1 percentage points overall. Only performance on recognizing colors was similar in level and gain to the Jump Start test; the control-group baseline Bracken score for colors was 56.4 percent and the increase 21.1 percentage points. The gains on the individual modules aside from colors were around 10 percentage points.

As was the case for the Jump Start test, the three estimates of the impact of the Jump Start program on the Bracken are similar in magnitude. Concentrating again on the last column, the Jump Start program is estimated to increase the overall test score by 2.2 percentage points (p-value=.02) more than the increase in the test score of the control group. Among the modules, the impact of the program is estimated to be 4.2

percentage points (p-value= .02) for Colors, 3.2 percentage points (p-value=.04) for Letters, 2.5 percentage points (p-value=.12) for Numbers, 1.2 percentage points (p-value=.32) for Sizes and 1.2 percentage points for Shapes (p-value=.36). The last two rows of table 7 also show the impact of the JSP on the national percentile score and percent of children delayed or very delayed. As seen, the control group's percentile score was essentially unchanged, while the percentile score of Jump Start participants increased by 3.5 points (p-value=.01). On the other hand, the impact of the program on percentage delayed or very delayed dropped by 4.1 percentage points (p-value=.16) on top of a 5.5 percentage point drop in the control group percentage.

5. The Mechanism

The RCT demonstrates that the children of parents trained to teach them the skills targeted by the JSP learned more of those skills than did children of parents not trained in the program. Moreover, relative to children of parents not participating in the JSP, the children of Jump Start parents acquired a greater level of skills not specifically targeted by the program (as measured by the Bracken). However, the RCT does not tell us how these results were achieved. To do that requires a model and additional data.

5.1. Model

To begin, we denote $T = 1$ if a family is in the treatment group (participates in the JSP) and $T = 0$ if the family is in the control group. Further, let K denote end-of-year knowledge and K_0 beginning-of-year knowledge. Assume the family maximizes a utility function subject to a knowledge production function and a budget constraint. Family i's utility is given by

$$U(C_i, K_i) = C_i + \gamma K_i \tag{1}$$

where C_i is household consumption and γ is the marginal utility of the child's end-of-year knowledge.

The end-of-year knowledge production function is assumed to satisfy the assumptions required for the value-added form (see Todd and Wolpin, 2006):

$$K_i = \kappa_0 + (1 + \delta)K_{0i} + \sum_{m=1}^{M} \beta_{im} \sqrt{X_{im}}, \qquad \beta_{im} \geq 0 \tag{2}$$

where the vector X_{im} comprises the full set of inputs in the production function of knowledge and the corresponding β_{im} are family-specific marginal product parameters. The β_{im} are allowed to depend on the treatment, namely:

$$\beta_{im} = \beta_{0im}(1 - T_i) + \beta_{1im}T_i, \qquad \text{for } m = 1, \dots, M. \tag{3}$$

The m^{th} input marginal product parameter is β_{0im} for a family that did not participate in the JSP and β_{1im} for a family that participated in the JSP. As captured in the model, the role of the JSP is to teach families how to help their children learn, that is, to increase their marginal products in utilizing inputs. Given that the JSP curriculum targets a specific subset of skills (see the difference between the Bracken and the JSP test), β_{1im} could be equal to β_{0im} for some inputs. On the other hand, there could be spillover effects of the program beyond the targeted skills if parents are induced by the program to expand the set of skills they work on with their child.

The family faces a one-period budget constraint: $C_i = Y_i - \sum_{m=1}^{M} p_m X_{im}$, where p_m is the price of input m.[8] Maximizing utility with respect to each input yields the input demand functions

$$X_{im} = \left(\frac{\gamma}{2p_m}\right)^2 (\beta_{im})^2, \qquad \text{for } m = 1, \dots, M. \tag{4}$$

Thus, if for a randomly drawn family i, $\beta_{1im} > \beta_{0im}$, the family will have a higher demand for input m as a participant in the JSP than as a non-participant. Note that the strong separability assumption in (2), the lack of complementarity or substitutability of inputs, implies that there are no cross-price effects in the demand for the inputs and that only the own-input marginal product parameter affects input demand.[9]

Substituting the input demand function (4) into the production function (2) provides the link between Jump Start participation and ending knowledge. Specifically,

$$K_i = \kappa_0 + (1 + \delta)K_{0i} + \sum_{m=1}^{M} \eta_m (\beta_{im})^2, \tag{5}$$

where $\eta_m = \frac{\gamma}{2p_m}$.

We can use (5) to derive the average treatment effect (ATE) obtained from the RCT. Let $K_i(0)$ and $K_i(1)$ denote, respectively, the end-of-year knowledge without and with participation in the JSP. Note that due to randomization, $\mathrm{E}[K_i(0)] = \mathrm{E}(K_i|T_i = 0)$, $\mathrm{E}[K_i(1)] = \mathrm{E}(K_i|T_i = 1)$, and $\mathrm{E}(K_{i0}|T_i = 0) = \mathrm{E}(K_{i0}|T_i = 1)$. Thus, $\mathrm{ATE} \equiv \mathrm{E}[K_i(1)] - \mathrm{E}[K_i(0)]$ is:

$$\mathrm{ATE} = \sum_{m=1}^{M} \eta_m \{\mathrm{E}[(\beta_{1im})^2] - \mathrm{E}[(\beta_{0im})^2]\}. \tag{6}$$

Equation (6) corresponds to the parameter that would be estimated from an OLS regression of ending knowledge on the treatment dummy. The parameter itself is a composite function of the utility and

[8] An equivalent structure given the lack of price data is that inputs enter directly into the utility function in an additively linear form. Time inputs, in particular, should be thought of as entering the model in this way.
[9] In addition, strong separability in (1) implies that there is no income effect in the input demand function.

production function parameters (and input prices). Understanding the mechanism that drives the impact of the JSP requires identifying the underlying parameters governing input choices. In particular, the effect of the JSP could be driven by families caring a great deal about their child's knowledge, γ is large, while the impact of the JSP on the parental marginal product parameters is small, or vice-versa.

Below we provide identification and estimation results for this model. The analysis addresses the issue of whether the mechanism proposed in the model is empirically consistent with the evidence from the RCT, that is, whether the (mean) marginal product parameter for specific input(s) is greater for the JSP families than for the control group families and whether Jump Start families choose more of the inputs.

The most direct method to answer the first question would be to estimate the production function (2). To place (2) into an estimation framework, rewrite it as

$$K_i = \kappa_0 + (1+\delta)K_{0i} + \sum_{m=1}^{M} \bar{\beta}_m \sqrt{X_{im}} + \sum_{m=1}^{M} \omega_{im,} \qquad (7)$$

where, on substituting the input demand equations (4), we obtain:

$$\omega_{im} = (\beta_{im} - \bar{\beta}_m)\eta_{im}\beta_{im}.$$

The problem in estimating (7) is that $E\left(\sqrt{X_{im}}\,\omega_{im'}\right) \neq 0$ in general, even though the inputs only depend on their own marginal product parameters, and even if β_{im} and $\beta_{im'}$ are assumed orthogonal. The ols estimate of $\bar{\beta}_m$ will be biased; the sign of the bias is in general indeterminate.[10]

Although without the model we cannot address whether the JSP improved the marginal productivities of families, we can determine whether, as the model implies, greater levels of inputs are chosen by JSP families. At the time of the recruitment and again at the end of the program year, families who agreed to be in the study were administered a survey. The survey included questions concerning the extent to which family members engaged in learning activities with their children. Table 8 shows the impact of the JSP on parental inputs: days/week read to the child, having at least 10 children's books, helping the child learn to identify letter, colors, numbers and shapes.

As seen in the table, as with test scores, parental inputs increased in the control group between the baseline and the end of JSP year. On average, parents in the control group increased the number of days someone in the family read to the child from 3.0 to 4.1 and the percentage of families having 10 or more children's

[10] If β_{im} and $\beta_{im'}$ are independently distributed, $E\left(\sqrt{X_{im}}\,\omega_{im'}\right) = \eta_m \eta_{m'} \sigma^2_{m'} \bar{\beta}_m$ for $m \neq m'$ and $E\left(\sqrt{X_{im}}\,\omega_{im}\right) = \eta^2_m \sigma^2_m \bar{\beta}_m + \mathrm{E}[\left(\beta_{im} - \bar{\beta}_m\right)^3]$. Absent skewness in the β_{im} distribution, the bias in $\bar{\beta}_m$ will be positive for all m.

books increased from 41 percent to 49 percent. In terms of parental help with particular skills, between 80 and 90 percent of the parents were already helping their children learn the alphabet, colors or numbers at baseline and by the end of the year the parental help increased by between 8 and 10 percentage points. Somewhat fewer parents were helping their child learn shapes at baseline, two-thirds, but that figure increased to 83.3 percent at the end of the year.

As also shown in table 8, at the end of the program, JSP families spent more days reading to the child than did the control group families and were more likely to have 10 or more children's books. Specifically, relative to control group families, JSP families read to their child on average one-half-day more and 22 percent more families had 10 or more books. The latter impact is almost surely because the JSP provided families with a dozen books as part of the curriculum; indeed 72 percent of the JSP families reported having 10 or more books (about the same as the graduation rate). The JSP had essentially no impact on the other parental inputs that we measured. This result is not surprising given the very high baseline percentages of families already providing these inputs to their children.

That the JSP increases the number of days spent reading to the child is a first step in the validation of the mechanism. The natural next step would be to estimate the model parameters, particularly to determine the size of the effect of the JSP on the productivity parameter. Given the results in table 8, we consider only the single input, the number of days spent reading to the child. If there are no omitted inputs, or no variation in them, as was essentially the case for the inputs in table 8 (we discuss the role of children's books below), the production function would look like (7) with $m = 1$. If there are relevant omitted inputs, then, denoting the included input with $m = 1$, the production function (7) becomes

$$K_i = \kappa_0 + (1 + \delta)K_{0i} + \bar{\beta}_1\sqrt{X_{i1}} + \zeta_i, \tag{8}$$

where $\zeta_i = \sum_{m=2}^{M} \bar{\beta}_m\sqrt{X_{im}} + \sum_{m=1}^{M} \omega_{im}$.

As noted, JSP families were provided children's books as part of the curriculum. Children's books are not by themselves direct inputs because they are of value only when combined, at a minimum, with the child's time. Having more books presumably affects the marginal product associated with parental reading input. Thus, the impact of the JSP on β_{i1} incorporates the provision of books.

1.1. Identification

The identification analysis, based on model moments, makes use of the input demand function (including additive measurement error) and the reduced form of the production function, specifically:

$$X_{i1} = \left(\frac{\gamma}{2p_1}\right)^2 (\beta_{i1})^2 + \varepsilon_{i1}, \tag{9}$$

$$K_i = \kappa_0 + (1+\delta)K_{0i} + \left(\frac{\gamma}{2p_1}\right)(\beta_{i1})^2 + u_i, \tag{10}$$

where ε_{i1} is measurement error, and $u_i = \sum_{m=2}^{M} \eta_m (\beta_{im})^2 + \upsilon_i$, where υ_i is a mean zero shock that arises in the derivation of the value-added production function (see Todd and Wolpin (2003)).[11] In addition, ε_{i1} and υ_i are assumed independent of K_{0i} and β_{i1}, and of each other

It is possible to establish an identification result based only on first-moments. Given the form in which β_{i1} enters (9) and (10), it is convenient to let

$$\mathrm{E}[(\beta_i)^2]_T = \bar{\beta}_T^2 + \sigma_{\beta_T}^2 \equiv F_0(1-T) + F_1 T. \tag{11}$$

The first moments of X and K are:

$$\mathrm{E}(X_i|T_i) = \left(\frac{\gamma}{2p}\right)^2 [F_0(1-T_i) + F_1 T_i], \tag{12}$$

$$\mathrm{E}(K_i|T_i) = \kappa_0 + (1+\delta)K_{0i} + \left(\frac{\gamma}{2p}\right)[F_0(1-T_i) + F_1 T_i] + \mathrm{E}(u_i|T_i). \tag{13}$$

Given that assignment to treatment is random, treatment identity can be used as an instrumental variable for the input. There are two possible Wald estimators for the impact of the input, one based on X_i and another on $\sqrt{X_i}$. Under the assumption that $E(u_i|T_i = 1) = \mathrm{E}(u_i|T_i = 0)$, that is, no other inputs were affected by the JSP, as was seen to be the case for the inputs in table 8, the elements of the Wald estimator based on X_i are:

$$\frac{\Delta \mathrm{E}(X_i|T_i)}{\Delta T_i} = \left(\frac{\gamma}{2p}\right)^2 [F_1 - F_0] \text{ and } \frac{\Delta \mathrm{E}(K_i|T_i)}{\Delta T_i} = \left(\frac{\gamma}{2p}\right)[F_1 - F_0] \tag{14}$$

The Wald (IV) estimator is thus

$$\frac{\frac{\Delta \mathrm{E}(K_i|T_i)}{\Delta T_i}}{\frac{\Delta \mathrm{E}(X_i|T_i)}{\Delta T_i}} = \left(\frac{\gamma}{2p}\right)^{-1} \tag{15}$$

Interestingly, this Wald estimator is not a function of the distribution of the marginal productivity parameter. Instead, as is obvious from (15), we can identify γ/p. In addition, from (12) for each treatment

[11] Specifically, $\upsilon_i = \tau_i - (1+\delta)\tau_{0i}$, where τ_i is the test-score measurement error associated with K_i and τ_{0i} the measurement error associated with K_{0i}.

group, we can also identify F_0 and F_1. The Wald estimator based on $\sqrt{X_i}$ provides the usual LATE estimator. However, because the measurement error is modeled as additive in X_i, and not in $\sqrt{X_i}$, it does not have a simple form. [12]

The rest of the identification argument relies on second moments. It also requires that either there be no omitted inputs as is the usual assumption in estimating test-score production functions or that the productivity parameters of all omitted inputs be homogeneous across families.[13] Although it may be possible to identify higher-order moments of the productivity parameter distribution, to carry out the exercise and in the empirical implementation we take a second-order approximation, letting

$$\mathrm{E}[(\beta_i)^4]_T = \bar{\beta}_T^4 + 6\bar{\beta}_T^2\sigma_{\beta_T}^2 \equiv G_0(1-T) + G_1 T.\text{[14]} \tag{17}$$

With that notation, the second moments for the treatment and control groups are:

$$\mathrm{Var}(X_i|T_i) = \left(\frac{\gamma}{2p}\right)^4 \{G_T - F_T^2\} + \sigma_\varepsilon^2 \tag{18}$$

$$\mathrm{Var}(K_i|T_i) = (1+\delta)^2\sigma_{K_0}^2 + \left(\frac{\gamma}{2p}\right)^2 \{G_T - F_T^2\} + 2(1+\delta)\left(\frac{\gamma}{2p}\right)\mathrm{Cov}\left(K_{0i}, {\beta_i}^2\right)_T + \sigma_u^2 \tag{19}$$

$$\mathrm{Cov}(X_i, K_{0i}|T_i) = \left(\frac{\gamma}{2p}\right)^2 \mathrm{Cov}\left(K_{0i}, {\beta_i}^2\right)_T \tag{20}$$

$$\mathrm{Cov}(X_i, K_i|T_i) = (1+\delta)\left(\frac{\gamma}{2p}\right)^2 \mathrm{Cov}\left(K_{0i}, {\beta_i}^2\right)_T + \left(\frac{\gamma}{2p}\right)^3 \{G_T - F_T^2\} \tag{21}$$

$$\mathrm{Cov}(K_i, K_{0i}|T_i) = (1+\delta)\sigma_{K_0}^2 + \left(\frac{\gamma}{2p}\right)\mathrm{Cov}\left(K_{0i}, {\beta_i}^2\right)_T \tag{22}$$

Given that γ/p is identified from the Wald estimator, $\mathrm{Cov}\left(K_{0i}, {\beta_i}^2\right)_T$ is identified from (20). A non-zero value of this covariance could arise because "initial" knowledge directly affects the marginal product and/or

---

[12] Had the measurement error been additive in $\sqrt{X_i}$, the Wald estimator would be given by

$$\frac{\dfrac{\Delta \mathrm{E}(K_i|T_i)}{\Delta T_i}}{\dfrac{\Delta \mathrm{E}\left(\sqrt{X_i}|T_i\right)}{\Delta T_i}} = \frac{(F_1 - F_0)}{\bar{\beta}_1 - \bar{\beta}_0} \tag{16}$$

[13] In this latter case, the model implies that all families choose the same level of these omitted inputs. Their impact is reflected in the constant term κ_0.

[14] The parameters of the distribution that are identified without further assumption about the distribution are $\bar{\beta}_T^2 + \sigma_{\beta_T}^2$ and $\mathrm{E}({\beta_i}^4)$. The second-order approximation is akin to assuming that the distribution is fully described by its first two moments.

because the marginal product parameter that produced K_{0i} is correlated with the parameter that produced K_i. Given these results, δ is identified from (22) and $G_T - F_T^2$ is identified from (21), which implies that σ_ε^2 is identified from (18) and σ_u^2 from (19).

Having previously identified F_T from the first moments, G_T is also identified. Then, for each F_T and G_T pair, there are two equations and two unknowns, $\bar{\beta}_T^2$ and $\sigma_{\beta_T}^2$. The solution for $\bar{\beta}_T^2$ must satisfy the polynomial equation:

$$5\bar{\beta}_T^4 - 6F\bar{\beta}_T^2 + G = 0. \tag{23}$$

Assuming the roots are real, this equation can either have one or two positive real roots for $\bar{\beta}_T^2$. If there are two positive roots, then both provide a valid estimate of $\bar{\beta}_T$. However, in either of these cases, the condition that $F_T - \bar{\beta}_T^2 = \sigma_{\beta_T}^2$ is greater than zero must also be satisfied.

1.2. Estimation

Measuring Child Knowledge K and K_0

As discussed, children were administered two separate assessments, the curriculum-based JSP test and the Bracken. As illustrated in tables 6 and 7, the two assessments together cover 10 distinct skill domains. Presumably, parents, the actors in the model, care about each of those domains, though not necessarily equally. Identification of the model requires a single scalar measure of knowledge. Two issues arise: (i) how to measure performance within any domain and (ii) how to combine domains.

In the previous discussion, performance was measured by the child's raw score on the JSP and Bracken assessments, that is, the percent of questions answered correctly. That choice was made because the question that AISD wanted to answer was whether the children in the program had mastered the JSP curriculum (as taught to them by their parents). The content of the JSP assessment followed the program curriculum exactly. For example, the children were supposed to have acquired the ability to recognize ten colors. The assessment asked the children to identify exactly those ten colors. The same was true in each of the six domains (see table 6) covered in the curriculum. The Bracken raw score was used as the performance measure to be comparable to the JSP measure.

The raw score is only one of several possible performance measures. Others include the percentile score for the tested population or, in the case of the Bracken, also the percentile score based on a nationally representative sample, a standardized score (e.g., mean zero, unit standard deviation), or an age-equivalence score, that is, the mean age of those with the score obtained by a given child. Although there is no strong reason to prefer one of these measures over the others, we chose the age-equivalence measure in the estimation of the model because it would be easily interpretable by parents and a measure they could form preferences over (your child is age a, but knows as much as children who are age $a + b$) and because it is a measure relative to a reasonable notion of their reference group, that is, children from other families that also applied to be in the JSP. It is also a measure that has a cardinal scale that is invariant to any monotonic transformation of the scores.[15]

In combining performance over domains, we use a weighted average of the age-equivalence scores of the separate domains. We assume that parents weight more heavily (as a matter of preference) performance in the more difficult domains. The weights we choose are the control group baseline average proportion of wrong answers in each domain (see tables 6 and 7).[16] Those weights are used both for the baseline and end-of-year measure of knowledge.

The impact of the JSP is to increase this weighted performance score by .41 age-equivalent years, or by 4.9 months.[17,18] The equivalent treatment effect based on a weighted average of raw scores is 7.6 percentage points; the baseline weighted test score for the control group is 30.8 and the weighted end-of-year test score is 47.1, an increase of 16.3 percentage points.

1.1. Results

Estimation is by Generalized Empirical Likelihood (GEL).[19] There are eleven model parameters, five of which are assumed to be the same for each treatment group, κ_0, δ, γ/p, σ_ε^2, σ_u^2, and three of which differ by treatment group, $\bar{\beta}_T$, $\sigma_{\beta_T}^2$, and $\mathrm{Cov}\left(K_{0i}, {\beta_i}^2\right)$. There are 14 data moments matched to the theoretical moments given by equations (12), (13), and (18)-(22).

The results of the estimation are shown in table 9. As discussed, the model restrictions did not rule out the possibility of there being two productivity distributions that would fit the data moments equally well. As seen in table 9, that turns out to be the case. In one set of estimates (the first column in table 9), parents in both treatment and control groups differ markedly in their productivity; the coefficient of variation for each

[15] The age-equivalence score is calculated by inverting a regression of the raw score for each domain on chronological age using the control group sample. Separate regressions were estimated for the baseline and end-of-year JSP and Bracken scores. The mean age equivalence for the control group at the baseline and at the end-of-year is thus the average chronological age of the control group sample at each time.

[16] The weights are normalized to sum to one.

[17] The regression also includes the baseline weighted age equivalent score. The p-value for the treatment effect is less than .001.

[18] The treatment effect based on the unweighted (simple average) of the age-equivalent scores is .36 years, or 4.3 months.

[19] The estimation was carried out using the GEL routine in the R-package, ‘gmm’.

group is about 2. In the second set of estimates, families are essentially homogeneous; the coefficient of variation is only about .1.

Not only do the two estimates of the distribution differ, but the way in which the JSP impacted parent productivities also differs. The first set of estimates imply that while increasing mean productivity, the JSP also created additional heterogeneity in parental productivity. Tests of the equality of the mean and variance of the productivity distribution, shown in the table, across the treatment and control group imply that the differences are statistically significant. On the other hand, the second set of estimates imply that the JSP increased mean productivity, without changing the degree of productivity heterogeneity among parents, which remained small.

The model fits the data well. Neither the Lagrange multiplier test nor J-test reject the over-identifying restrictions. As implied by these tests, the difference between the actual and predicted fitted moments is small; the percentage difference relative to the data moments is smallest for the means, less than .5 percent, largest for the covariances, ranging from less than one percent to 7 percent with a single outlier being 18 percent, and in-between for the variances, less than 2.5 percent.

How important, in this population, is reading to one's child in affecting the acquisition of the skills in the domains covered by the JSP and Bracken assessments? It is possible to get a picture of the overall importance of the reading input in determining ending knowledge. Based on the estimated model, the proportion of the variance in ending knowledge accounted for by the reading input, that is by $\beta_i\sqrt{X_i}$ is 42 percent for the control group and 45 percent for the treatment group.[20] In addition, given that about 10 percent is accounted for by variation in beginning knowledge, under the model assumptions less than half of the variation in ending knowledge is due to omitted inputs. This result is more striking given that the model estimates imply that about 88 percent of the variance in the measured reading input is due to measurement error.[21] Thus, the actual standard deviation in the input is only .67 days per week in the control group and .75 in the treatment group (as opposed to 1.89 and 1.91 days for the input measure).

Our results show that the JSP increased the average value of $\beta_i\sqrt{X_i}$ by changing the distribution of β_i, and, as a response by parents to the increase in their productivity, by changing the level of the input chosen by the JSP families. As already noted, this response was substantial; parents increased the number of days per week they read to their child on average by a little over one-half a day. We also found, based on the estimated treatment effect, that the change in the productivities of families substantially increased their

[20] In this calculation, we ascribe the contribution of the covariance between beginning knowledge and the productivity parameter (see equation (18)), to the reading input.
[21] The input is measured as the average over the baseline and end-of-year number of days/week the parent read to the child.

children's knowledge as measured by the increase in their age equivalency (4.9 months). The model, which was not rejected by the data, provides a consistent interpretation of the mechanism behind the success of the JSP.

What conclusions might a researcher have drawn about the productivity of the parental input without the model? Standard practice would have been to estimate some version of the value-added production function under the assumption that the productivity parameter is the same across families. It is important to recognize that the additional assumptions of the model about the orthogonality between included and omitted inputs are also necessary to obtain consistent estimates of the production function estimated outside of the model.

Table 10 provides three alternative estimates of the production function. The first column shows an ols regression of the impact of parent reading time on the age-equivalent test score using the model's production function specification, allowing for separate effects of the input for the treatment and control group. The estimate is .182 for the control group and .399 for the treatment group. In contrast, the estimates from the model are 1.60 and 1.71 (based on the estimates in the second column of table 9). As shown in table 10, the p-value for the difference in the treatment and control group estimates is .069.

The second column shows the results from an ols regression that constrains the control and treatment group coefficients to be the same. In that regression, the treatment effect is assumed to operate only through parent reading time.[22] The impact of parent reading time on the age-equivalent test score is .336. The third column uses treatment identification as an instrument for parent reading time as a means of correcting for measurement error in the input.[23] The estimate is almost an order of magnitude greater than the ols estimate in the second column, consistent with the large estimated degree of measurement error in the input.[24]

Based on these results, one might conclude that the IV regression provides a good estimate of the mean of β. However, in the context of the model, this conclusion would be incorrect. As previously shown, the IV-regression does not estimate the productivity parameter, but rather the LATE parameter (16), which, in the case where there is heterogeneity in β, depends on the treatment and control group's mean and variance of β. Interestingly, the LATE parameter, in the case that β is homogenous across families, but the JSP improves parent productivity, reduces to the sum of the treatment and control group $\beta's$.[25]

2. Conclusion:

[22] We impose this constraint to be able to contrast the result to the IV regression in the third column. Allowing for a separate treatment effect reduces the input coefficient to .291 (.062).
[23] Given the assumption of homogeneity in the marginal product parameter, there is no other reason to use the treatment as an instrument.
[24] The measurement error attenuation factor estimated from the model is .12.
[25] This result depends on the measurement error being additive in the square root of the input.

This paper reports the results of a study that evaluated the Jump Start Program designed by the Alief Independent School District located in the Houston area. We found that the program produced modest impacts on the acquisition of the skills targeted by the program curriculum, about .25 of a standard deviation on raw test scores. We also found some evidence of spillovers to content not directly covered by the curriculum as measured by the Bracken Assessment of School Readiness scale. The program increased parental investments as measured by the frequency parents read to their children.

To go beyond simply measuring treatment effects, we built and estimated a model of parent-child interactions that featured individual-specific heterogeneity in the marginal productivity of investments. We developed a constructive identification analysis and implemented estimation based on it. The model estimates implied that the program impact on children's knowledge is due to an increase in the marginal productivity of parental time spent reading to their child, which increases parental efficiency and, in response, induces an increase in the input itself, that is, in the frequency with which parents choose to read to the child.

The JSP trained parents to engage more effectively with their children on a very limited set of skills. Because these skills are perceived to be foundational for Pre-K success by AISD teachers, it is possible that the children whose parents participated in the program will have medium- and long-term benefits. It is also possible that these foundational skills are easily learned in the first weeks of the Pre-K program and any short-term advantage conferred by participation in the program will be short-lived. To answer these questions, future work will analyze medium- and long-term impacts of the program on performance as the children in the Jump Start study progress through school.

References

Attanasio, Orazio , Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina. 2015. Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia. In *Working Papers*, edited by NBER.

Baker, Amy J.L., Chaya S. Piotrkowski, and Jeanne Brooks-Gunn. 1998. "The Effects of the Home Instruction Program for Preschool Youngsters (HIPPY) on Children's School Performance at the End of the Program and One Year Later." *Early Childhood Research Quarterly* 13 (4):571-588.

Baker, Amy J.L., Chaya S. Piotrkowski, and Jeanne Brooks-Gunn. 1999. "The Home Instruction Program for Preschool Youngsters (HIPPY)." *Home Visiting: Recent Program Evaluations* 9 (1):116-133.

Bassok, Daphna, Jenna E. Finch, RayHyuck Lee, Sean F. Reardon, and Jane Waldfogel. 2016. "Socioeconomic Gaps in Early Childhood Experiences: 1998 to 2010." *AERA Open* 2 (3):1-22.

Bond, Timothy N., and Kevin Lang. 2013. "The Evolution of the Black-White Test Score Gap in Grades K-3: The Fragility of Results." *Review of Economics and Statistics* 95 (5):1468-1479.

Cunha, Flavio, Marsha Gerdes, and Snejana Nihtianova. 2019. Language Environment and Maternal Expectations: An Evaluation of the LENA Start Program. edited by Rice University.

Cunha, Flavio, and James J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2):31-47.

Cunha, Flavio, James J. Heckman, and Susanne Schennach, M. . 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3):883-931.

Florens, Jean-Pierre, James J. Heckman, Costas Meghir, and Edward Vytlacil. 2008. "Identification of Treatment Effects Using Control Functions in Models with Continuous Endogenous Treatment and Heterogeneous Effects." *Econometrica* 72 (1):1191-1206.

Heckman, James J. , Margaret L. Holland, Kevin K. Makino, Rodrigo Pinto, and Maria Rosales-Rueda. 2017. An Analysis of the Memphis Nurse-Family Partnership Program. In *Working Papers*, edited by NBER.

Kalil, Ariel 2014. "Inequality Begins at Home: The Role of Parenting in the Diverging Destinies of Rich and Poor Children." In *Families in an Era of Increasing Inequality: Diverging Destinies*, edited by Paul R. Amato, Alan Booth, Susan M. McHale and Jennifer Van Hook, 63-82. New York: Springer.

Kalil, Ariel, Rebecca Ryan, and Michael Corey. 2012. "Diverging destinies: maternal education and the developmental gradient in time with children." *Demography* 49 (4):1361-1383. doi: 10.1007/s13524-012-0129-5.

Love, John M. , Ellen Eliason Kisker, Christine Ross, Helen Raikes, Jill Constantine, Kimberly Boller, Jeanne Brooks-Gunn, Rachel Chazan-Cohen, Louisa Banks Tarullo, Christy Brady-Smith, Allison Sidle Fuligni, Peter Z. Schochet, Diane Paulsell, and Cheri Vogel. 2005. "The Effectiveness of Early Head Start for 3-Year-Old Children and Their Parents: Lessons for Policy and Programs." *Developmental Psychology* 41 (6):885-901.

Putnam, Robert D. 2016. *Our Kids: The American Dream in Crisis*. New York: Simon and Schuster.

Reardon, Sean F., and Ximena A. Portilla. 2016. "Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry." *AERA Open* 2 (3):1-18. doi: 10.1177/2332858416657343.

St. Pierre, Robert, Anne Ricciuti, Fumiyo Tao, Cindy Creps, Janet Swartz, Wang Lee, Amanda Parsad, and Tracy Rimdzius. 2003. Third National Even Start Evaluation: Program Impacts and Implications for Improvement. Department of Education, Washington, DC.

Suskind, Dana L. , Kristin R. Leffel, Eileen Graf, Marc W. Hernandez, Elizabeth A. Gunderson, Shannon G. Sapolich, Elizabeth Suskind, Lindsey Leininger, Susan Goldin-Meadow, and Susan C. Levine. 2015. "A parent-directed language intervention for children of low socioeconomic status: a randomized controlled pilot study." *Journal of Child Language* 43 (2):366-406.

Todd, Petra E., and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal* 113 (485):F3-F33.

Todd, Petra E., and Kenneth I. Wolpin. 2207. "Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps." *Journal of Human Capital* 1 (1):91-113.

Appendix A

Alief Jump Start Curriculum Weeks 1-22

Month/Week	Materials
October Week 1	-Play Dough -No, David! book -CD - "Name Song" -Laminated Sentence Strip -Laminated Construction Mat -Expo Marker
October Week 2	-Play Dough -The Pigeon Needs a Bath! Book -CD - "Name Song" -CD - "Brush a Brush" -Hand Pump Soap -Tooth Brush and Tooth Paste -Laminated Sentence Strip -Laminated Construction Mat -Expo Marker - Small Pack of Tissues
October Group Week	-CD - "Name Song" -CD - "Brush a Brush" -Pre Made sentence strips with each child's name on it- laminated -Hand Sanitizer -No, David! book
November *Week* 4	-CD - Color Songs "Red" -CD - "The Color Song" -Brown Bear, Brown Bear Book -Construction paper- red -Glue bottles -Leaf outline -Apple outline
November Week 5	-CD - Color Son·gs "Blue",. -CD - "The Color Song" -Brown Bear, Brown Bear Book -Construction paper- blue -Glue bottles -Butterfly outline -Bird Outline
November Group Week	-CD - Color Songs "Red" "Blue" -Brown Bear, Brown Bear book -Sentence Strip with names pre- written in blue and red -Red and Blue beads -Red and Blue Markers

December Week 7	-CD - Color Songs "Green" -CD - "The Color Song" -Scissors - Chicka Chicka, Boom Boom book -Green construction paper -Green, red, blue strips of construction paper
December Group	CD -"Color Song "Green" -Sentence strip pre-written with name in green. -Dr. Suess's ABC: An Amazing Alphabet Book!" -Tree outline -Green, blue and red sticker dots -Scissors
December Week 9	-CD - Color Songs "Green" -CD - "The Color Song" -Scissors - Dr. Suess's ABC: An Amazing Alphabet book -Green construction paper -Half sheet of white construction paper -Green sticker dots -Frog Outline -Alligator Outline -Glue -Play dough
January Week 10	-CD - Color Songs "Yellow" -CD - " Five Little Fish" CD - "Five Little Monkeys" -Large Crayons -Rainbow Fish book -White copy paper -3 white die-cut fish -Yellow construction paper -Ocean scene outline -Cotton balls -Glue
January Week 11	-CD - Color Songs "Black" - Track 10 -CD - " Five Little Fish" - Track 8 CD- "Five Little Monkeys" -Track 9 -Large Crayons -Ten Black Dots book -Dot to Dot paper -Spider to Dot paper -Cutting Line Paper -Spider Web -5 Black Spiders

January Group Week	-CD - Color Song "Black" - Track 10 -CD - Color Song "Yellow" - Track 7 -CD - "Five Little Monkeys" - Track 9 -Sentence strip pre-written with name in black. -Ten Black Dots book -Bee outline -Yellow construction paper -Scissors -Glue -Wax paper wings -Black crayon -Black sticker dots
February Week 13	-CD - Color Songs "Purple" - Track 11 -Harold and the Pur12le Cray:on book -Tongs -Porn Porns -Sorting mat -Clear plastic cup
February Week 14	-CD - Color Songs "Brown" - Track 12 -The Three Bears book -Goldilock cutting paper -Play Dough -Counting Mat -Die Cut Counting Bears (16 total) -Sorting Mat
February Group Week	-CD - Color Songs "Purple" - Track 11 -CD - Color Songs "Brown" - Track 12 -Sentence strip pre-written with name in purple. -The Three Bears book -Goldilocks Rhyme -Popsicle Sticks -Bear outline -Goldilocks Die Cut -Purple Marker
March Week 16	-CD - Color Songs "Orange" - Track 13 -CD - "Head, Shoulder, Knees and Toes" Track 14 -From Head to Toe Book -Simon Says poster
March Week 17	-CD - Color Songs "Orange" - Track 13 -CD - "Head, Shoulder, Knees and Toes" Track 14 - If You're Ha1212Y: and You Know It book -Circle Outline

i

	-Glue -Scissors -Crayons -Black and Blue Dots -Copy paper -Play Dough -Counting Mat -Brown Construction Paper
March Group Week	-CD - Color Songs "Orange" - Track 13 -CD - "Head, Shoulder, Knees and Toes" Track 14 -From Head To Toe book -Pre Written sentence strip with child's name on it in orange.
April Week 19	-CD - Color Songs "Pink" - Track 15 -Beads -Yarn -Cut Up Straws -Pipe Cleaners -The Three Little <u>Pigs</u> book -Buttons -Pink Construction Paper -Crayons - Pig Outline
April Week 20	-CD - Color Songs "Pink" - Track 15 -CD - "Five Little Monkeys" - Track 9 - The Three Little Pigs book -Spinning Tops -Play Coins -Crayons -Pig Coloring Page -Cookie Cutters -Sentence Strip
April Group Week	-CD - Color Songs "Pink" - Track 15 -CD - "Head, Shoulder, Knees and Toes" -Track 14 -The Three Little Pigs book -Pre Written sentence strip with child's name on it in pink. -Name Puzzle (Pre made on sentence strips) -Pink Construction Paper -Glue
MayWeek22	-CD - Color Songs "White"-Track 16 -CD - "The Color Song" - Track 6 -"Mary Had a Little Lamb" Rhyme Sheet -"Humpty Dumpty" Rhyme Sheet

-

	-Lamb Outline -Glue -Cotton Balls -Tweezers -Counting Mat -Red Construction Paper -Crayons -Scissors -Egg Outline -Black Crayon -Humpty Dumpty's Wall Outline

Table 1
Sample Sizes

	In Jump Start (Treatment Group)	Not In Jump Start (Control Group)	Total
Year 1			
In Lottery			
Number	232	229	461
Percent	50.3	49.7	100.0
In Study			
Baseline			
Number	211	205	416
Percent	50.7	49.3	100.0
Percent of Lottery	90.9	89.5	90.2
End of Year			
Number	170	127	297
Percent	57.2	42.8	100.0
Percent of Lottery	73.3	55.4	64.4
Percent of Baseline	80.5	62.0	71.3
Year 2			
In Lottery			
Number	199	190	389
Percent	51.2	48.8	100.0
In Study			
Baseline			
Number	184	178	362
Percent	50.8	49.2	100.0
Percent of Lottery	92.5	93.7	93.1
End of Year			
Number	160	138	298
Percent	53.7	46.3	100.0
Percent of Lottery	80.4	72.6	76.6
Percent of Baseline	86.9	77.5	82.3
Year 3			
In Lottery			
Number	190	176	366
Percent	51.9	48.1	100.0
In Study			
Baseline			
Number	179	169	348
Percent	50.8	49.2	100.0
Percent of Lottery	94.2	96.0	95.1
End of Year			
Number	157	138	295
Percent	53.2	46.8	100.0
Percent of Lottery	82.6	78.4	80.6
Percent of Baseline	87.7	81.7	84.8

Table 2
Comparison of Treatment and Controls: Baseline Study Sample

Variable	Mean	Treatment-Control Difference	p-Value of Difference
Blocking Variables			
Jump Start Baseline Test Score (%)			
Year 1	51.4	1.0	.70
Year 2	53.5	0.2	.93
Year 3	48.4	0.01	.99
Pooled Years	51.2	0.4	.78
Child Age at Test (Years)			
Year 1	3.58	0.01	.84
Year 2	3.52	0.01	.80
Year 3	3.46	-0.01	.94
Pooled Years	3.52	0.01	.88
Non-Blocking Variables			
Parent Years of Schooling			
Year 1	12.2	0.16	.64
Year 2	12.1	0.11	.76
Year 3	12.3	0.09	.81
Pooled Years	12.2	0.12	.55
Parent Hispanic			
Year 1	57.0	-2.0	.66
Year 2	57.4	-3.1	.54
Year 3	58.8	0.3	.95
Pooled Years	57.7	-1.7	.56
Parent Age (years)			
Year 1	32.5	-0.12	.84
Year 2	32.4	0.84	.28
Year 3	32.3	0.20	.78
Pooled Years	32.5	0.28	.48
Number of Children in HH			
Year 1	2.87	0.07	.58
Year 2	2.81	0.09	.47
Year 3	2.82	0.17	.21
Pooled Years	2.83	0.11	.15
Percent Receive Food Stamps (%)			
Year 1	51.2	-3.5	.49
Year 2	45.0	-6.4	.22
Year 3	46.0	2.0	.72
Pooled Years	47.6	-2.8	.35
Bracken Baseline Test Score			
Year 1	25.2	1.2	.54
Year 2	23.5	0.6	.77
Year 3	22.8	1.0	.60
Pooled Years	23.9	0.9	.42

Table 3
Comparison of Treatment and Controls: Post-Program Study Sample

Variable	Mean	Treatment-Control Difference	p-Value of Difference
Blocking Variables			
Jump Start Baseline Test Score (%)			
Year 1	50.1	1.0	.74
Year 2	53.9	0.5	.86
Year 3	48.8	1.3	.67
Pooled Years	51.0	0.9	.59
Chile Age at Test (Years)			
Year 1	3.58	0.02	.52
Year 2	3.51	-0.01	.91
Year 3	3.46	-0.03	.36
Pooled Years	3.52	-0.01	.96
Non-Blocking Variables			
Parent Years of Schooling			
Year 1	12.4	0.02	.96
Year 2	12.1	0.46	.27
Year 3	12.3	0.45	.26
Pooled Years	12.2	0.32	.18
Parent Hispanic			
Year 1	60.6	1.3	.82
Year 2	60.0	-6.1	.29
Year 3	61.0	-3.8	.51
Pooled Years	60.5	-2.9	.39
Parent Age (years)			
Year 1	32.6	-0.23	.78
Year 2	32.7	0.73	.41
Year 3	32.4	-0.15	.84
Pooled Years	32.6	0.12	.79
Number of Children in HH			
Year 1	2.96	0.15	.30
Year 2	2.80	-0.03	.84
Year 3	2.85	0.09	.50
Pooled Years	2.87	0.07	.36
Percent Receive Food Stamps (%)			
Year 1	50.5	0.9	.88
Year 2	42.4	-8.0	.17
Year 3	45.4	2.3	.69
Pooled Years	46.1	-1.5	.66
Bracken Baseline Test Score			
Year 1	23.6	0.5	.83
Year 2	25.1	3.2	.17
Year 3	22.9	0.8	.70
Pooled Years	23.9	1.5	.24

Table 4
Jump Start Program Completion By Year

	Year 1	Year 2	Year 3
Mean Weeks Completed	13.5	17.4	16.1
Pct. Completing			
0 weeks	10.4	2.0	8.4
1 – 5 Weeks	15.7	10.6	11.1
6 - 15 Weeks	16.6	13.1	8.9
Pct. Complete Program (16+ weeks)	57.2	74.4	71.6

Table 5
Baseline Test Scores:
Study Participants in Pre- and Post-Test Sample (All Years)

	Year 1	Year 2	Year 3	All Years
Jump Start Test				
Mean Test Score (%)	50.1	53.9	48.8	51.0
Standard Deviation (%)	26.4	25.0	25.6	25.7
% Test Score< 25.0%	20.6	16.4	21.7	19.7
% Test Score>= 75.0%	20.6	22.8	17.6	20.3
Bracken Test				
Mean Test Score (%)	23.6	25.1	22.9	23.9
Standard Deviation (%)	19.1	20.2	17.7	19.0
Mean National Percentile	25.6	28.8	26.9	27.1
Percent Very Delayed or Delayed	52.4	48.7	51.2	50.7

Table 6
Impact of Jump Start Program on the Jump Start Test – All Years
Alternative Estimators[1]

Test Module (no. of questions) : [control group baseline test score, control group end-of-year test score)	End-of-Year Difference	Diff-in-Diff (1)	Diff-in-Diff (2)[2]
All Modules (26): [50.4, 68.7]	7.4	6.5	7.0
	(1.4)	(1.5)	(1.2)
Colors (10): [58.4, 74.7]	3.7	2.7	3.2
	(1.9	(2.1)	(1.6)
Fine Motor Skills (8): [56.9, 75.7]	3.8	4.6	4.0
	(1.5)	(2.1)	(1.4)
Counting (2): [40.3, 62.7]	3.8	0.7	2.8
	(2.9)	(3.4)	(2.7)
Sorting (1): [42.9, 67.2]	6.5	5.9	6.4
	(3.1)	(4.2)	(3.0)
Name Recognition (1): [24.1, 47.4]	22.4	21.0	22.2
	(3.2)	(4.0)	(3.2)
Book Handling (4): [31.2,48.5]	22.1	19.2	21.5
	(2.5)	(2.9)	(2.4)

1. Robust standard error in parentheses.
2. Based on regression of end-of-year test score on treatment dummy and baseline test score.

Table 7
Impact of Jump Start Program on the Bracken Test - All Years
Alternative Estimators[1]

Test Module (no. of questions): [control group baseline test score, control group end-of-year test score)	End-of-Year Difference	Diff-in-Diff (1)	Diff-in-Diff (2)[2]
All Modules (85): [23.0, 34.1]	3.2 (1.3)	1.6 (1.1)	2.2 (1.0)
Colors (10): [56.4, 77.4]	5.3 (2.2)	2.9 (2.4)	4.2 (1.8)
Letters (8): [20.2, 28.2]	4.6 (2.1)	2.6 (1.7)	3.2 (1.6)
Numbers (18): [11.9, 20.6]	4.2 (1.9)	1.8 (1.7)	2.5 (1.6)
Sizes (22): [20.3, 29.1]	1.1 (1.3)	1.4 (1.5)	1.2 (1.2)
Shapes (20): [21.5, 34.9]	2.1 (1.5)	0.4 (1.5)	1.2 (1.3)
Percentile Score[3] (85): [25.9, 26.0]	4.8 (1.8)	2.6 (1.6)	3.5 (1.4)
Pct. Very Delayed or Delayed[3] (85): [51.0, 46.5]	-4.3 (3.4)	-4.0 (3,5)	-4.2 (3.0)

1. Robust standard errors in parentheses.
2. Based on regression of end-of-year test score on treatment dummy and baseline test score.
3. National Norm by Age in Months

Table 8

Impact of Jump Start Program on Parental Inputs - All Years

Input [control group baseline, control group end of year]	Diff-in-Diff [1]
Number of Days/Week Read to Child: [3.0, 4.1]	0.56 (0.14)
Has 10 or More Children’s Books: [.41, .49]	0.22 (0.03)
Helps Child Learn Alphabet: [80.8, 87.6]	0.02 (0.03)
Helps Child Learn Colors: [89.6, 99.0]	0.02 (0.02)
Helps Child Learn Numbers: [88.3, 95.8]	-0.01 (0.02)
Help Child Learn Shapes: [67.6, 83.3]	-0.04 (0.03)

1. Based on regression of end-of-year value of variable on treatment dummy and baseline value of variable; robust standard errors in parentheses.

Table 9

Model Parameter Values (standard error in parentheses)

Parameters			
$\bar{\beta}_0$	0.73 (0.18)		1.60 (0.40)
$\bar{\beta}_1$	0.78 (0.29)		1.71 (0.42)
σ_{β_0}	1.42 (0.36)		0.15 (0.03)
σ_{β_1}	1.53 (0.37)		0.16 (0.03)
γ/p		2.36 (0.58)	
$\mathrm{cov}\left(K_{0i}, {\beta_i}^2\right)_1$		0.27 (0.16)	
$\mathrm{cov}\left(K_{0i}, {\beta_i}^2\right)_1$		0.43 (0.21)	
σ_u		0.82 (0.06)	
σ_ε		1.76 (0.05)	
κ_0		0.15 (0.66)	
δ		-0.75 (0.06)	
Tests of Over-Identifying Restrictions (Statistic, p-value)			
Lagrange Multiplier		1.64, 0.65	
J-Statistic		1.66, 0.65	
Tests of Equality: D-Statistic; $\chi^2_{.01}$=6.63			
$\bar{\beta}_0 = \bar{\beta}_1$	19.3		19.2
$\sigma^2_{\beta_0} = \sigma^2_{\beta_0}$	20.1		0.08
$\mathrm{cov}\left(K_{0i}, {\beta_i}^2\right)_0 = \mathrm{cov}\left(K_{0i}, {\beta_i}^2\right)_1$		1.47	

Table 10
Production Function Estimates

	Ols	Ols	IV
$\sqrt{X}$		.336	2.91
		(.062)	(.816)
$(1-T)*\sqrt{X}$	.182[1]		
	(.091)		
$T*\sqrt{X}$	.399[1]		
	(.079)	-	-
K_0	.475	.479	.288
	(.026)	(.026)	(.071)
T	-.041		
	(.233)		
Constant	2.08	1.99	-2.18
	(.183)	(.133)	(1.35)
R^2	.384	.359	

1. p-value for test of coefficient equality = .069