Linking Incongruent Data Sources: A Case Study of a Summer Library Program

Jennifer A. Boden, Karin Chang, and Meghan Ecker-Lyster
Kansas City Area Education Research Consortium, University of Kansas, USA

Introduction

A wealth of empirical evidence demonstrates that the educational achievement gap is exacerbated by economic inequality. Compared with children from advantaged backgrounds, children from low-income families are less likely to score as proficient on school reading assessments,[1] graduate from high school,[2] and attend college (Goldrick-Rab & Pfeffer, 2009; Bailey & Dynarski, 2011).[3] Although there are many underlying causes of income-based disparities, low-income children are more vulnerable to summer learning loss than their wealthier peers.[4]

To better understand summer learning loss, the "faucet theory" suggests that opportunities to learn and access educational resources are "turned on" for all children during the academic year and are accessed equally; however, the school resources are "turned off" during the summer months when school is not in session.[5] As a result of the faucet being shut off during the summer months, low-income students whose families cannot afford to provide supplemental educational resources and learning opportunities are put at a distinct disadvantage compared to their affluent peers. On average, children from middle-income homes have access to approximately twelve books per child, whereas children from low-income families have access to about one book per child.[6]

To combat summer learning loss, a number of districts and communities have implemented summer reading programs. Local public libraries are a key player in the delivery of summer reading programs, as this is one community establishment that offers all children, regardless of income level, equal access to learning opportunities (e.g., books). Over 95% of libraries in the US offer a summer reading program.[7] Research has found promising evidence regarding the success of public library summer reading programs on enhancing student reading outcomes.[8]

The literature on summer learning loss and summer reading programs address major aspects of the efficacy of these interventions on student outcomes; however, the literature leaves important questions around measurement largely unanswered. Many evaluation studies utilize the same standardized reading assessment, such as the Scholastic Reading Inventory (SRI), to measure the impact of library reading programs on student outcomes.[9] Few studies explicitly explore the methods necessary to successfully leverage extant reading data across multiple school districts, which often contain myriad reading assessments.

Study Purpose and Research Questions

A major challenge to evaluating the impact of a library summer reading program (SRP) is accessing the necessary data to adequately measure program success. Many programs do not have the resources nor time to administer a specific reading assessment to participants. A cost-effective alternative is to obtain reading data from the local school districts that send students to the library's reading program. However, school districts choose benchmark assessments individually and there is considerable variability across schools, grade levels, and content areas in terms of which benchmark assessment is used. Thus, working with multiple districts, many of whom choose different assessments, poses some methodological challenges.

The purpose of this study was to develop a data-driven research method that could be used to measure and evaluate the impact of SRP. The primary research question that guided the evaluation study was: Do students who participate in an SRP experience comparable levels of summer learning loss compared to similar students who did not participate in the program?

Data

The SRP that was evaluated in this study was developed by the Mid-Continent Public Library which serves the Kansas City Metropolitan Area (KCMA). Data for this study were derived from two primary sources: the SRP program records and extant academic data from participating school districts. Data for this analysis corresponded to three time periods. Pre-intervention data were denoted **Spring** and came from the spring of academic year 2016–17. Post-intervention data were denoted **Fall** and came from the fall of academic year 2017–18. Intervention data were denoted **Participation** and corresponded to the intervention which occurred during the summer of 2017.

Approximately 11,000 students from 21 public school districts as well as various private and charter schools across the KCMA participated in the 2017 SRP. Data were solicited from 16 of the 21 public school districts, and data sharing agreements were ultimately executed with eight. The 16 districts were chosen primarily due to size. The data sharing agreement specified that **Spring** and **Fall** datasets be submitted as separate rather than aggregate files. This specification necessitated additional time to be set aside to create a matched **Spring-Fall** dataset. Commentary regarding the rationale behind this decision and its applicability to areas of high student mobility may be see in Appendix A. The distribution of participants across districts and by participation status may be seen in Figure 1.

Spring and **Fall** data were administrative data provided by the eight participating public-school districts. **Spring** and **Fall** data elements consisted of identifiers such as name, birth date, state identification numbers, demographic information, and benchmark reading scores. **Participation** data were self-reported by participants and then provided to us by MCPL. The data corresponded to students who participated in the 2017 SRP. **Participation** variables consisted of name, birthdate, school, and grade. Note that **Participation** data did not include numeric identifiers; this created challenges, as discussed later in the paper, when joining the various datasets. A full list of requested elements may be seen in the sample Memorandum of Understanding located in Appendix B.

Methods

The analytic sample for this study consisted of a **treatment** group, defined as first through fifth grade students who participated in the 2017 SRP, and a **control** group of similarly-aged students who did not participate. Figure 2 presents the grade distribution of all participants in the SRP as reported by the parents of these participants or by the participants themselves to the MCPL staff. Notice that a substantial percentage of participants did not report grade information. Due to the higher likelihood of inaccuracy in these data, observations were dropped based on district data rather than **Participation**. More specifically, student records were dropped if they could not be linked to district data at all or if, once linked to the district data, students appeared to be outside the range of first through fifth grades. Students who were in kindergarten or younger during spring 2016–17 were dropped from the analytic sample due to lack of assessment measures. Students older than fifth grade were dropped due to small sample sizes.

Linking Data Sources

The decision to match **Spring** with **Participation** data rather than **Fall** with **Participation** was arbitrary, but the decision to match district data to **Participation** before matching district to district was intentional. As evaluators, our primary responsibility was reporting on the efficacy of the program, but due to the annual replication of this study, it was equally important that we assess the quality of data collection related to participation and make recommendations accordingly. **Participation** data, particularly the identification variables such as name and date of birth, were prone to errors, missing information, and, as discussed in Appendix C, informal variations of names. We wanted to know how much the quality of these data influenced our ability to match **Participation** to district data. By matching **Participation** to **Spring** first, we could determine the percentage of SRP participants that were identifiable in the district data through name-matching. This percentage provides insight into the quality of the **Participation** data and the related collection process.

Standardizing Assessment Scores

The raw data for this evaluation reported student assessment scores across five different reading measures: Lexiles, Fountas & Pinnell (F&P), Curriculum-Based Measure of Oral Reading (RCBM), Rausch Unit Scale (RIT), and STAR Reading (STAR) Scores. Because these assessment scores were based on varying scales, we standardized each to get scores that are comparable across all districts and grades. Student records containing only F&P scores were dropped from the analytic sample. This is because F&P scores are non-numeric and currently there is no agreed-upon method by which to quantify these data. While this only affected one of the eight districts, it was the largest of the participating districts and dropped a substantial number of observations. The equation used to standardize the remaining assessments was:

$$Z_i = \frac{Assessment\ Score_i - Sample\ Average}{Sample\ Standard\ Deviation}$$

Sample averages and standard deviations were functions of assessment type and grade. For example, the z-score calculation for a third-grade student taking the RCBM assessment utilized the sample average and standard deviation of the third-grade students across the eight districts that took the RCBM assessment. However, when assessment companies publish norming sample means and standard deviations, these statistics were used in place of our sample-based statistics.

The benefit of standardizing assessment scores was two-fold. First, it allowed us to keep observations of students who took a different assessment in the fall than what he or she took in the spring. This was particularly important because it was not uncommon for districts to change assessments from one academic year to the next, and, with a highly mobile student population, it was also likely that students would move from one district to another that took a different assessment. The second benefit of standardizing assessment scores was that scores became comparable across all grades and assessment types. Because assessment scores were then interpreted as the number of standard deviations that a given score was away from the corresponding mean, the downside of standardizing was that it could be difficult for practitioners and lay audiences to understand the findings. The ease in which we could convert z-scores into something meaningful depended on the audience (e.g., parents, practitioners, administrators, etc.), but was necessary because of the inconsistency of assessment types.

Results
Nearest Neighbor Matching

One problem with basic regressions is that involvement in the intervention is not randomly assigned. If students in the **treatment** and **control** groups had similar characteristics resulting in an equal likelihood of entering treatment, this would not be an issue. However, **treatment** and **control** groups look somewhat dissimilar in terms of observable characteristics. Figures 3 and 4 illustrate some of these differences.

Treatment group students were more likely to be white (61 percent compared to 53 percent), and less likely to be male (47 percent compared to 50 percent), black (17 percent compared to 23 percent), or Hispanic (12 percent compared to 14 percent). Asian and other racial groups were represented similarly in both the **treatment** and **control** groups.

It also appeared that **treatment** students were also less likely to receive free or reduced lunch, but it was difficult to determine whether that was actually the case. The Community Eligibility Provision (CEP) allows districts that are 40% FRL or higher to designate all students as FRL. One of the districts in our sample used this provision and this may have altered the distribution of this variable across treatment designations. **Treatment** group students were much more likely to be enrolled in summer school (68 percent compared to 33 percent). Although summer school attendance may have been another important factor in determining the effect of the SRP intervention on student reading outcomes, this variable is missing in a substantial number of cases.

Because of these differences in demographic characteristics and the fact that assignment to the intervention is not random, we opted to create a matched sample.

Creating the Matched Sample

A nearest-neighbor matching algorithm was used to create the **control** group. The nearest-neighbor algorithm estimates the counterfactual for each observation by identifying one or more students who are similar in terms of a collection of designated observable characteristics. More specifically, the comparison group for a treated student will be made up of one or more students who are untreated but have similar observable characteristics. **Treatment** and **control** students matched exactly on the district that they attended in the spring, otherwise, **treatment** and **control** students were match based on a weighted function of gender, race, English language learner status, and spring test score. The reason that we included the spring test score as a matching element is that we wanted to ensure that students were starting out at a similar level of proficiency. The logical argument behind this idea is that if we look at two students who are similar in all ways except their exposure to the intervention, then the more likely it is that the intervention is the cause for any differences in the outcome variable.

Once the match was made, the estimate of the counterfactual outcome variable was calculated. The average difference between all students' actual and estimated counterfactual outcomes is called the average treatment effect (ATE).

Table 1 presents our nearest-neighbor estimates of the ATE of the SRP intervention on student reading outcomes. Compared to similar students, SRP participants had better fall reading outcomes following the intervention. The estimated average treatment effect was 1.660 (p = 0.082), meaning that reading outcomes would be 1.660 standard deviations higher when all students participate in the SRP compared to when no students participate. The output also indicates that ties-in-distance caused at least one observation to be matched with five observations. What this means is that the nearest-neighbor matching algorithm identified all equally good observations and averaged the associated outcomes to calculate the counterfactual in an effort to reduce bias.

Matching **treatment** students to **control** students with the same summer school status increased the ATE to approximately 5.3 (p = 0.002). This estimate means that, when summer school is taken into account, if **treatment** and **control** students alike participated in the SRP, the reading assessment outcome would, on average, be 5.3 standard deviations higher than it would be otherwise. This estimate is significant at the 1 percent level. Including FRL status with (6.7, p = 0.000) and without summer school status generates estimates that are also significant at the 1 percent level (2.7, p = 0.015).

Conclusion

Leveraging secondary data sources is a cost-effective approach to evaluating community programs, such as a summer library program. However, complexity is introduced when data is solicited from multiple sources (e.g., school districts, public library), which pose many challenges for linking and summarizing outcomes. The purpose of this study was to highlight a data-driven, research method that can be used to measure and evaluate program outcomes that rely on data from numerous sources. This case study highlights a methodological approach that was used to evaluate the impact of a summer library program on reducing summer learning loss. Results indicated that the SRP was an effective intervention for reducing the impact of summer learning loss for participants.

1. Jencks & Phillips, 1998; Reardon, 2011.
2. Losen, 2004; Rumberger, 2011.
3. Losen, "Graduation rate accountability" and Rumberger, *Dropping Out*.
4. Cooper, et al., "The effects of summer vacation" and Entwisle, "Summer learning and home environment"
5. Entwisle, et al., "Summer learning and home environment."
6. Celano and Nueman, "When schools close."
7. NCES, *Services and resources*.
8. Roman and Fiore, "Do public library" and Shin and Krashen, *Summer reading program*.
9. Roman and Fiore, "Do public library."

References

Bailey, M. J., and S. M. Dynarski. "Gains and Gaps: Changing Inequality in U.S. College Entry and Completion." *National Bureau of Economic Research, Working Paper No. 17633.* Cambridge, MA, December 2011, https://www.nber.org/papers/w17633.

Cooper, H., B. Nye, K. Charlton, J. Lindsay, and S. Greathouse, S. "The effects of summer vacation on achievement test scores: A narrative and meta-analytic review." *Review of Educational Research* 66, no. 3 (September 1, 1996): 227–268. https://doi.org/10.3102%2F00346543066003227.

Entwisle, D. R., K. L. Alexander, and L. S. Olson. "Summer learning and home environment." In *A nation at risk: Preserving public education as an engine for social mobility*. New York, NY: Century Foundation Press, 9–30.

Celano, D., and S. B. Neuman. "When schools close, the knowledge gap grows." *Phi Delta Kappan 90* (2008): 256–263.

Goldrick-Rab, S., and F. T. Pfeffer. "Beyond Access: Explaining Socioeconomic Differences in College Transfer." *Sociology of Education* 82 (2009): 101–25.

Kim, J. "Summer reading and the ethnic achievement gap." *Journal of Education for Students Placed at Risk* 9 (2004): 169–178.

Kim, J. S. and D. M. Quinn. "The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and home interventions." *Review of Educational Research* 83 (2013): 386–431.

Jencks, C., and M. Phillips. *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution, 1998.

Los Angeles County Public Library Foundation. *Evaluation of the public library summer reading: "Books and Beyond . . . Take Me to Your Reader." Final report.* Los Angeles, CA: Public Library Foundation, 2001.

Losen, D. J. "Graduation rate accountability under the No Child Left Behind Act and the desperate impact on students of color." In *Dropouts in America: Confronting the Graduation Crisis*, 41–56. Edited by Gary Orfield. Cambridge, MA: Harvard Education Press, 2004.

National Center for Education Statistics (NCES). *Services and resources for children and young adults in public libraries*. 2014. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=95357.

Shin, F. H., and S. D. Krashen. *Summer reading program and evidence*. Boston, MA: Pearson, 2008.

Reardon, S. F. "The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations." In *Withier Opportunity? Rising Inequality, Schools, and Children's Life*

Chances, 91–115. Edited by Greg J. Duncan and Richard J. Murnane. New York, NY: Russell Sage Foundation, 2011.

Roman, S., and C. D. Fiore. "Do public library summer reading programs close the achievement gap?" *Children and Libraries* 8, no. 3 (2010): 27–35.

Rumberger, R. W. ***Dropping Out: Why Students Drop Out of High School and What Can Be Done About It.*** Cambridge, MA: Harvard University Press, 2011.

Tables
Table 1: ATE Estimates

	Without Summer or FRL	With Summer (exact match)	With FRL	With FRL and Summer
Average Treatment Effect	1.660*	5.288***	2.650**	6.676***
Standard Error	(0.955)	(1.695)	(1.091)	(1.855)
With Summer (exact)	12,039	10,256	9,540	7,756

Figures
Figure 1: Distribution of SRP participants across school districts

Percent SRP Participants By District and Participation Status (N = 10,934)

Figure 2: Grade distribution of SRP Participation data
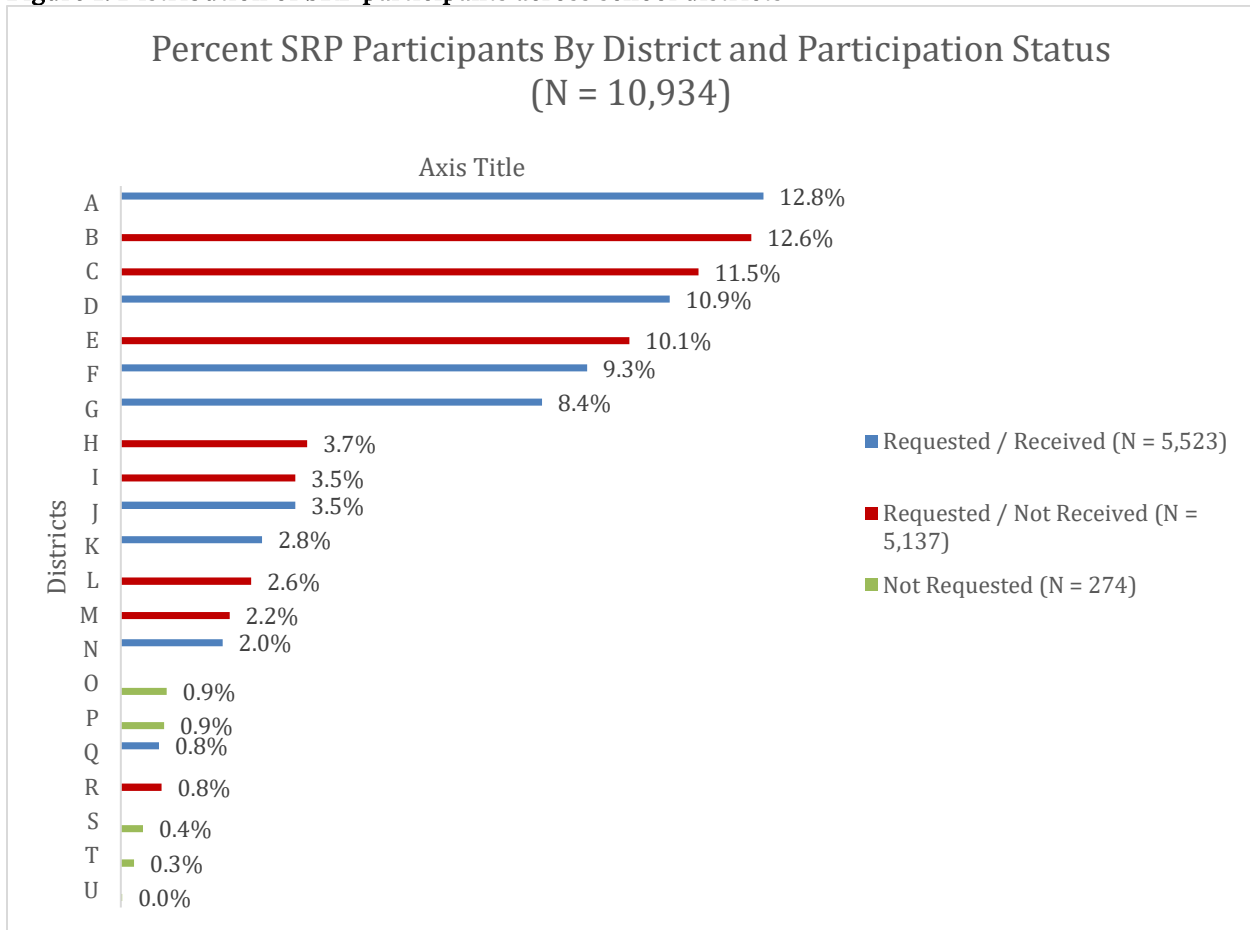

Grade Distribution of SLP Participants
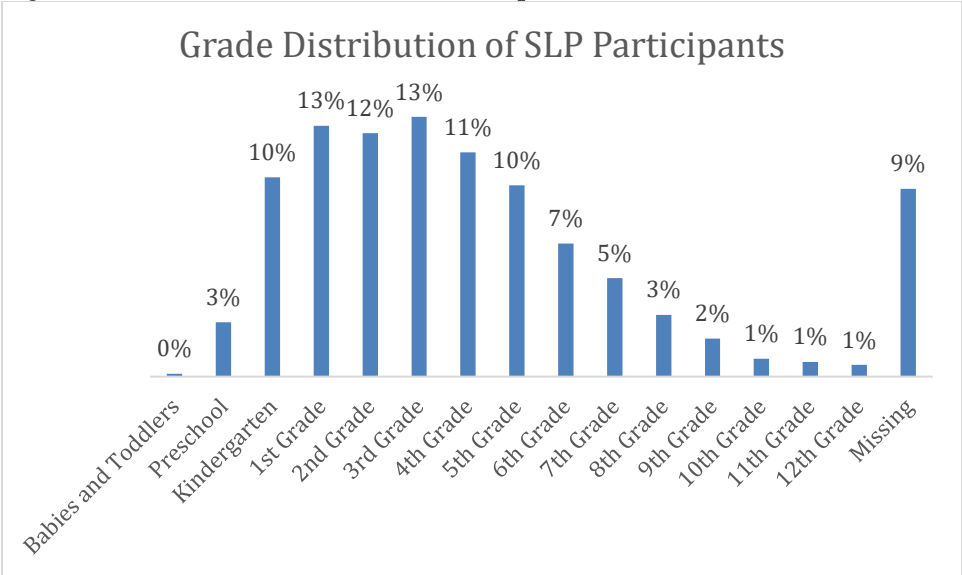
Figure 3: Demographic Characteristics by Treatment Status

Race, Ethnicity, and Gender Distributions by Treatment Status

- White: Control 53%, Treatment 61%
- Black: Control 23%, Treatment 17%
- Hispanic: Control 14%, Treatment 12%
- Male: Control 50%, Treatment 47%

Control / Treatment
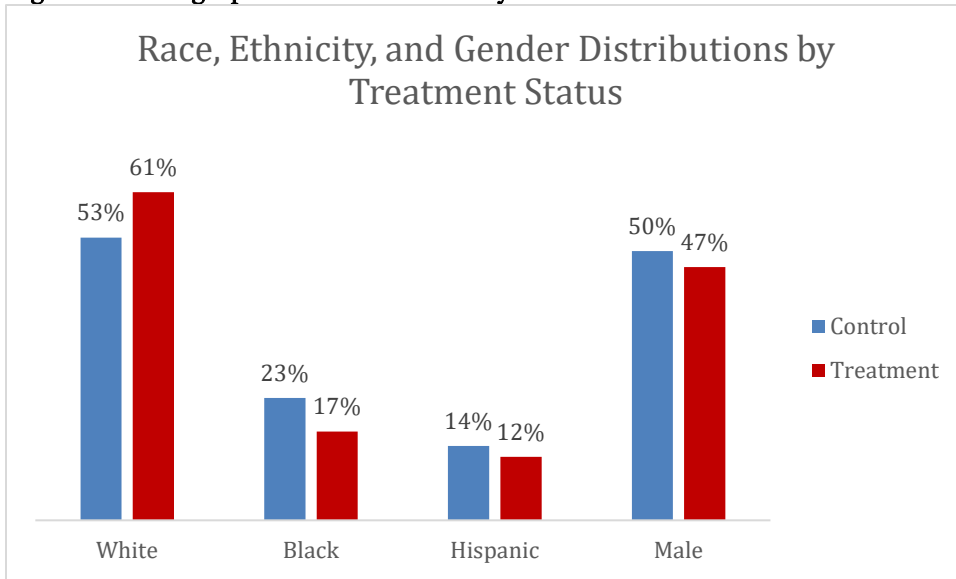
Figure 4: Demographic Characteristics by Treatment Status

Selected Characteristics by Treatment Status

- Summer: Control 33%, Treatment 68%
- FRL: Control 48%, Treatment 45%
- SPED: Control 14%, Treatment 13%
- ELL: Control 5%, Treatment 3%

Control / Treatment
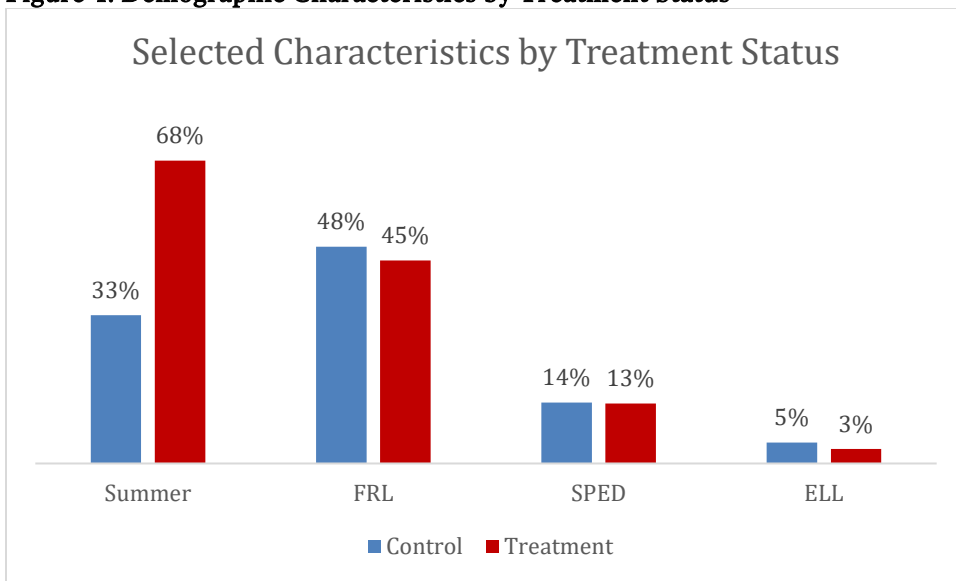
Appendix A: Separate versus Aggregate District Data

In the early years of the evaluation process, districts were asked to submit spring and fall data as a single dataset rather than separate files. While not needing to merge spring and fall datasets saved considerable time, there was one major disadvantage. Some of the participating districts have highly mobile student populations, and when submitting the data for analysis, districts often deleted any incomplete student records. In other words, when a student was not enrolled for both spring and fall semesters, districts often dropped the student from the district data submission. Incomplete records are usually dropped during analysis, so under typical circumstances it would not matter whether the district submitted the incomplete student records. Since we work with a number of districts in close proximity to each other, we have an opportunity to identify missing student data across multiple districts. For example, if a student attended District A during the spring and then transferred to District B during the fall, if District B is one of our participating districts, then we can maintain that student's record in the analytic sample by simply matching his partial record from District A to his partial record from District B. Although districts would have been willing to resubmit the data to include incomplete student records, we tried to avoid this whenever possible. Because the continued success of this project hinges on district participation, it is important to minimize the effort required by district staff.

Because we received two data files, **_Spring_** and **_Fall_**, from each district for the summer 2017 evaluation, we recoded variables for individual districts and then created master spring and master fall datasets which contained all student records from all districts for the corresponding semester. The master **_Spring_** dataset was matched with MCPL **_Participation_** data, and then the resulting data were matched with the master **_Fall_** dataset.

Appendix B: Example Memorandum of Understanding

Memorandum of Understanding between <u>SCHOOL DISTRICT</u>, MID-CONTINENT PUBLIC LIBRARY and the UNIVERSITY OF KANSAS CENTER FOR RESEARCH, INC.

WHEREAS, <u>School District</u>, the Mid-Continent Public Library and the University of Kansas Center for Research, Inc. (KUCR) on behalf of the Kansas City Area Education Research Consortium (KC-AERC) wish to create an independent, non-partisan vehicle of the very highest quality to evaluate the effect of the summer reading program efforts and contribute to basic scholarly research on public schools and educational programming; and

WHEREAS, in order to advance these goals, it is necessary to create a digital data archive consisting of longitudinal data that have been fully cleaned, integrated, and documented; and

WHEREAS, data on student characteristics, student academic performance, and school characteristics are necessary to address the foreseeable research questions of the Consortium and the public it serves; and

WHEREAS, to achieve these purposes, the Consortium will release standardized data to a broad public while protecting the individual-level confidentiality.

Now therefore, the parties agree as follows:

1. <u>School District</u> will appoint a data liaison to coordinate this work (at <u>School District</u>) who will facilitate access to the data, and arrange for the staff resources necessary to create all data files to be provided to KUCR.

2. The Mid-Continent Public Library will appoint a liaison to coordinate this work, facilitate access to the data, and arrange for the staff resources necessary to conduct the project. The library will support data technician(s) on KUCR staff or as consultants, as necessary and commensurate with the scale and scope of actual data transferred.

3. Once this MOU has been fully executed by all parties, KUCR will provide resources to clean, organize, match, and manage all data files provided for the project. KUCR will design and execute a methodology for analyzing the data.

4. School District will, through a data liaison, or other representative, resolve in a timely fashion through discussions with the Executive Director or other staff of KUCR any questions that arise concerning the content, timing, or other aspects of the data transfer.

5. School District will provide the staff resources needed to assist, in a timely fashion, with the design and documentation of the data, and create or supply the extract files for KUCR from the administrative systems of the School District.

6. School District and Mid-Continent Public Library will provide KUCR with the data needed to sustain the mission of evaluating efforts of the reading program. School District will provide extracts containing the following types of student information for every student enrolled in pre-kindergarten through 12th grade for the school years 2016–2017 and 2017–2018:

Spring, Academic Year 2016–2017	Fall, Academic Year 2017–2018
MOSIS (or KIDS) student identification number (scrambled using an agreed-upon algorithm)	MOSIS (or KIDS) student identification number (scrambled using an agreed-upon algorithm)
Student first name	Student first name
Student middle name	Student middle name
Student last name	Student last name
Student date of birth	Student date of birth
Student grade level	Student grade level
Student race	
Student gender	
Student free lunch status	
Student reduced lunch status	
Student special education	
Student English Language Learner (ELL) status	
Student benchmark reading and math assessment scores (e.g., i-Ready, STAR, AIMSweb)	Student benchmark reading and math assessment scores (e.g., i-Ready, STAR, AIMSweb)
Indicate if student attended district summer school in 2017	

7. It is anticipated that modifications to this MOU will be issued for future requested data elements for the following school years: 2018–2019, 2019–2020, and 2020–2021. Any modification of this MOU shall be in writing and shall be signed by both parties.

8. <u>School District</u> will provide KUCR with the data described above for students in pre-kindergarten through 12th grade attending all schools for which <u>School District</u> maintains data, including elementary, alternative schools, collaborative programs and special education programs. The feasible beginning date (i.e., historical and longitudinal character) of the data will be determined in consultation with <u>School District</u>.

9. KUCR will preserve the confidentiality of all personally identifiable information about all individual students obtained pursuant to the Memorandum of Understanding in accordance with applicable law, including the Federal Social Security Act, the Family Educational Rights and Privacy Act and any regulations promulgated there under. All studies will be conducted in a manner that does not permit personal identification of parents, teachers and students by persons other than required for research activities undertaken by representatives of KUCR. As such, KUCR will not disclose any such information to any persons except as authorized by law and upon formal approval of <u>School District</u> and will include results in aggregate or in some other non-personally identifiable form. KUCR assures all researchers who are given access to data with individual-level identifiers provided pursuant to the Memorandum of Understanding will have undergone appropriate training.

10. KUCR will subject all research initiated under this Memorandum of Understanding to review and approval by KUCR's Human Research Protection Program, as applicable. KUCR may publish results, analysis, or other information developed as a result of any research based on the data made available under this agreement only in summary or aggregate form, ensuring that no personally-identifiable information is disclosed.

11. KUCR will create a standardized series of data files for broader public release. Standardized data is defined as aggregate school-level data or individual-level data that has been stripped of individual-level identifiers and cannot generate any possible multivariate analysis combining data fields that would yield less than five records per any data cell. Standardized data files will be reviewed for considerations of accuracy and privacy by <u>School District</u> prior to public release, if a public release is planned.

12. The agreement between <u>School District</u>, the Mid-Continent Public Library, and KUCR is effective as of the date of the last signature and shall continue 60 days after the evaluation agreement authorized by MCPL and KUCR ends. It is anticipated that the evaluation agreement will continue through 02/28/2021, unless terminated earlier by either party. Either party may terminate this agreement provided written notification is received by the other party 30 days prior to the proposed termination date.

By signing below, the official certifies that he or she has the authority to bind the organization to the terms of this Understanding and that the organization has the capability to undertake the commitments in this Understanding.

Appendix C: Difficulties Related to Name-matching

As previously noted, **Participation** data do not include a numeric identifier, so the matching process between **Participation** and **Spring** datasets relied primarily on names. Name-matching these two datasets was a time-intensive endeavor for three main reasons: use of informal name in the **Participation** dataset, inconsistent spelling of names, and missing information. The biggest challenge of the three is the use of informal names. District data provide formal or legal names while **Participation** data are typically more informal (i.e., nicknames). For example, the name "Theodore" in **Spring** might be listed as "Ted," "Teddy," or "Theo" in **Participation**. Although these variations on the name "Theodore" make algorithmic matching difficult, it is still possible to join this student's data from one dataset to the other because these variations are commonplace and widely known. Where this process may become impossible is when informal names are unrelated to the formal name. This is particularly prevalent in districts with high immigrant populations where some students may register under anglicized names. For example, a student legally named "Fan" may decide to sign up for the SRP using the name "Sam." Because these names are phonetically unrelated it may be impossible to match this student's district record with his **Participation** record. It is possible to circumvent this problem by using additional information such as grade, date of birth, and last name, but even with this additional information, students are often unmatchable.

To reduce the amount of time needed to match **Spring** and **Participation** files, we utilized a user-written command in STATA called "matchit" (from Julio Raffo). Matchit measures the distance between two text strings and produces a similarity score for the pairing. If two text strings match exactly then the similarity score is equal to 1, and all other matches result in a similarity score less than 1. Exact matches were automatically kept, and all other potential matches were considered individually. Once **Spring** and **Participation** were matched, the resulting dataset was then matched to **Fall**. This part of the matching process was straightforward in that **Spring** and **Fall** datasets contained numeric identifiers upon which we could connect student records from one semester to another.

Appendix D: Duplicate Student Observations with Mismatched Outcome Data

Once **Spring** and **Participation** are matched, the resulting dataset is matched to **Fall**. The matching process is straightforward in that **Spring** and **Fall** datasets contain numeric identifiers upon which we can connect student records from one semester to another.

One thing to be aware of during this stage of the data preparation is that students, who should appear only once in the fully matched set, sometimes appear more than once. This happens occasionally when students move from one district to another, so student information may exist in both the previous district as well as the current district. For example, a student lives in District A and takes the District A assessment. During that same spring, the student moves into District B, and District B uses a different assessment than District A. District B has the student take their assessment during that same spring. During the fall, the student moves to District C. In the fully matched dataset, the student will have two records. One record will be District A in the spring with District C in the fall, and the other will be District B in the spring and District C in the fall. We have no reason to conclude that one record is more correct than the other, but each student may only appear in the analytic sample once. Which observation pair do we choose? If we decide to always choose the observation pair with the higher spring score, or likewise with the lower spring score, then we may introduce bias into our estimate of the effect of SRP on student achievement. Although we do not observe this phenomenon very often, it is still important to avoid introducing bias if possible, so we randomly select one observation for students with more than one record.