Levers for Fostering Quality Improvement in Louisiana Early Childhood Education Programs

Justin Brian Doromal, MEd
Frank Batten School of Leadership and Public Policy
University of Virginia

May 2018

Page intentionally left blank.

Levers for Fostering Quality Improvements in Louisiana Early Childhood Education Programs

Justin Brian Doromal, MEd

Frank Batten School of Leadership and Public Policy
University of Virginia

Prepared for:
Jenna Conway
Assistant Superintendent – Early Childhood
Louisiana Department of Education

May 2018

Disclaimer: The author conducted this study as part of the program of professional education at the Frank Batten School of Leadership and Public Policy, University of Virginia. This paper is submitted in partial fulfillment of the course requirements for the Master of Public Policy degree. The judgments and conclusions are solely those of the author, and are not necessarily endorsed by the Batten School, by the University of Virginia, or by any other entity.

Honor Pledge: On my honor as a student, I have neither given nor received unauthorized aid on this report.

Acknowledgments

I am sincerely grateful to Jenna Conway at the Louisiana Department of Education for her enthusiasm and willingness for me to pursue this project, and for providing me with the resources to prepare this report.

I am also incredibly grateful to Daphna Bassok for connecting me with the Louisiana Department of Education, and for providing me with invaluable mentorship these past several years.

I would like to thank James Wyckoff for his helpful feedback during our seminar, for thoughtfully reviewing my drafts, and for pushing me to think critically about my work.

I am appreciative of my colleagues for helping me think through important pieces of this report during seminar, and for challenging me to think differently about various policy issues. I am especially thankful for Katharine Sadowski, for her numerous suggestions and insights, as well as Walter Herring, Anandita Krishnamachari, Emily Wiseman, and Kylie Anglin for their friendship and moral support.

Finally, I am indebted to the love and support of my family and friends who unceasingly inspire and motivate me.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ACF	Administration for Children and Families
ACS	American Community Survey (2012-2016 5-year estimates)
Act 3	Early Childhood Education Act
CCAP	Childcare Assistance Program
CDA	Child Development Associate
CLASS	Classroom Assessment Scoring System
ECE	Early Childhood Education
LDE	Louisiana Department of Education
MMCI	Making the Most of Classroom Interactions
MTP	My Teaching Partner
NSECD	Nonpublic School Early Childhood Development (Program)
PD	Professional Development
QRIS	Quality Rating and Improvement System(s)

Page intentionally left blank.

EXECUTIVE SUMMARY

Rigorous evidence demonstrates that the skills children develop in *high-quality* early childhood education (ECE) programs can influence their educational and labor market outcomes, particularly when children are exposed to rich and engaging interactions with their teachers and peers (Phillips et al., 2017). Yet in Louisiana, nearly *one third* of ECE programs exhibited quality that inadequately supports children's learning and development and prepares them for success in school.

Louisiana has invested substantially toward developing a unified system for monitoring and ensuring quality across all publicly-funded ECE programs and is keenly interested in how to support programs' efforts toward improvement. Despite this vision for unified quality standards, child care centers and Head Start programs continue to be disproportionately lower quality, while school-based preschools compose nearly all the state's highest rated programs on quality.

The first aim of this report was to identify and evaluate the strategies program leaders were able to employ—particularly those in non-school-based settings—in response to state-defined goals for improvement. Toward this end, I analyze survey responses from 482 ECE program leaders in Louisiana, linking financial data on programs' quality improvement efforts, to scores derived from the Classroom Assessment Scoring System (CLASS), a observational measure of classroom quality that is tied to the state's quality rating system.

I evaluate three types of quality improvement decisions: investments in *classroom resources*, such as the use of state-approved, high-quality curricula; investments toward *teacher preparation*, such as credentialing and certification; and investments in *professional development for teachers*, such as training and coaching. I find that decisions that help prepare and support teachers—for example, coaching models like *Making the Most of Classroom Interactions* (MMCI), CLASS observer trainings, and ensuring teachers hold professional certification in ECE—show the greatest promise for effectively fostering quality improvement, even above and beyond other determinants of quality such as program and parish characteristics.

The second aim of this report was to understand policy options by which Louisiana programs could <u>efficiently</u> and <u>equitably</u> foster quality improvement. I propose the following policy levers:

1. Implement effective group coaching models, such as MMCI;
2. Increase teacher participation in *CLASS PreK observer reliability training*; or
3. Ensure more teachers ($\geq$50%) have earned the Child Development Associate credential.

I recommend that program leaders consider adoption of group coaching models like MMCI. My analysis suggests this lever is not only the most cost-effective of the three options examined, it is also the most politically feasible and is positively received by participating teachers. Given that MMCI has relatively high start-up costs associated with training of coaches, and that MMCI is currently not widely accessed, the Louisiana Department of Education should consider targeting resources to ensure equitable access to these effective improvement levers.

BACKGROUND AND CONTEXT

Problem Statement

Thirty-one percent of early childhood education (ECE) programs in Louisiana are not of sufficient quality to support children's learning and development, and only half of Louisiana children enter kindergarten ready to succeed. Although the state has invested substantially in a unified system that helps program leaders understand on which dimensions of quality to improve, program leaders may still be challenged in identifying and employing levers that effectively and efficiently improve their quality.

Policy Context

In 2010, an estimated 87% of publicly-funded early childhood education (ECE) programs across the nation were not of high enough quality to support children's learning and development (Burchinal, Vandergrift, Pianta, & Mashburn, 2010). This finding was striking, and called on policymakers to improve the quality of these settings given the importance of the early years for kindergarten readiness and later life success (Phillips et al., 2017).

Louisiana was receptive to this call: In 2012 the state passed the Early Childhood Education Act (Act 3), which overhauled Louisiana's ECE system with the explicit goal of building a cohesive system that supported improvements in school readiness. Act 3 was a multifaceted, multi-year effort (see Figure 1 below) that aimed to ensure that all ECE programs were held to the same standard of quality, so that family circumstances did not determine the quality of children's education. These investments were unprecedented and reflected Louisiana's growing commitment to ECE.

Act 3 achieved several key provisions. First, to address a fragmented early childhood system that inadequately prepared children for school success, oversight of child care and other non-school-based ECE programs was shifted to the Louisiana Department of Education (LDE), offering greater cohesion with the state's existing public preschool program and the K-12 education system. Indeed, the ECE system now serves all children through age five, including infants and toddlers. Second, across all ECE programs receiving public dollars, the LDE implemented a statewide kindergarten entry assessment, *Teaching Strategies GOLD*, that allowed teachers to maintain portfolios for all children and uniformly assess their development on a number of school readiness domains. Teachers could use this tool to track children's learning over the year and share portfolios with kindergarten teachers and parents to facilitate their kindergarten transition.

Third, to ensure programs were providing children with the rich experiences needed to promote early learning, the LDE developed and implemented a quality rating and improvement system

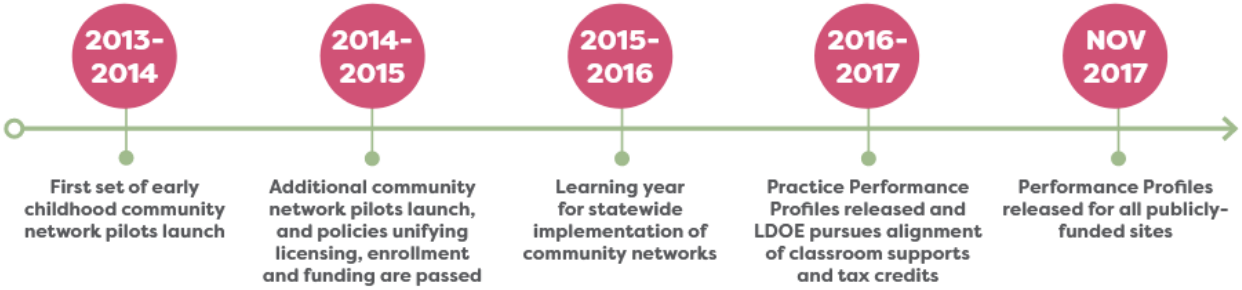

Figure 1. Timeline of Louisiana's multi-year effort to unify its early childhood system and improve kindergarten readiness.

(QRIS) for measuring and assessing the quality of ECE programs. QRIS, which are accountability systems that measure, incentivize, and improve quality at scale (Zellman & Perlman, 2010), are a widely adopted and promising lever for supporting quality improvement: nearly every state implements a QRIS, and 40 states have a statewide QRIS. Under the QRIS approach, state-level decisions about how to define and measure quality, as well as how to incentivize and support quality improvement, 'cascade' down and influence the practice of ECE program leaders and teachers. Importantly, QRIS does not prescribe specific actions that programs should take toward improvement.

Finally, the LDE established "community networks," which locally oversaw all ECE programs and were responsible for making key decisions to ensure community needs were being met, including ensuring adequate access to high-quality ECE programs and engaging in quality improvement activities. Indeed, although the LDE enacted many of the reforms central in efforts to build cohesion and improvement, the 65 community networks, often parishes, were ultimately the agents of this change.

Louisiana's Quality Rating and Improvement System

Louisiana's QRIS is the primary vehicle through which quality in ECE is measured and *improvements* in quality are incentivized. QRIS generally aim to improve quality through five core components: (1) well-designed quality standards, (2) quality monitoring and accountability, (3) support for program improvement, (4) financial incentives linked to quality, and (5) easily accessible quality information provided to parents.

States typically operationalize "quality" using a complex set of factors including both structural features of the classroom (e.g. teacher-child ratios) and classroom observations. However, recent critiques of QRIS have posited that states should streamline QRIS to focus on the most relevant aspects of quality (Cannon, Zellman, Karoly, & Schwartz, 2017). Louisiana's QRIS responds to these criticisms by focusing only on aspects of quality shown by the research to support children's development: rich and engaging teacher-child interactions (Araujo, Carneiro, Cruz-Aguayo, & Schady, 2016; Burchinal et al., 2000, 2010; Early et al., 2007; Mashburn et al., 2008; Pianta, Barnett, Burchinal, & Thornburg, 2009).

In fact, while Louisiana does publish individualized, easy-to-understand *Performance Profiles* that summarize program characteristics (see Appendix A), the quality of classroom processes as measured by the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008) is the *only* determinant of Louisiana programs' quality rating. The CLASS, which is a well-validated and widely used research tool for characterizing the quality of classroom interactions (Hamre, Hatfield, Pianta, & Jamil, 2014; Pianta, La Paro, et al., 2008), uses a 7-point scale to assess classrooms on a number of domains important for children's learning, described in Figure 2 at right. These domains broadly

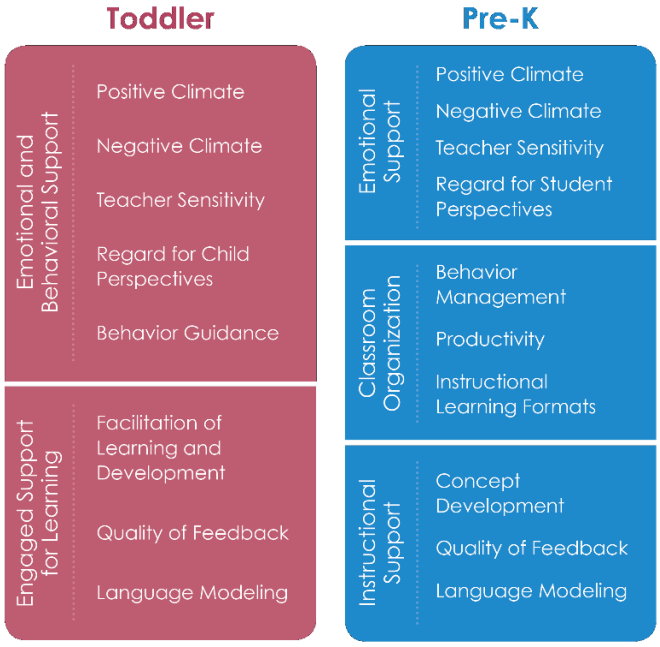

Figure 2. CLASS domains for quality classroom interactions.

capture characteristics of the teacher and classroom that support warm and engaging interactions, and they are conceptualized differently by age group to acknowledge children's different developmental needs across the early years.

As part of its QRIS, Louisiana has conducted *at least* two observations per year using the CLASS, in *every* classroom within *all* ECE programs serving publicly-funded toddlers or preschoolers (ages 2-5), since the 2015-16 school year. Classrooms' scores are then aggregated to the program-level to determine a program's quality rating. Programs with average CLASS scores between 6-7 are labeled "Excellent," those with scores between 4.50-5.99 are "Proficient," 3-4.49 corresponds to "Approaching Proficient," and those with average scores below 3 are "Unsatisfactory." These score brackets are public knowledge, such that providers are able to access their underlying CLASS scores and can reference their performance relative to quality rating cutoffs.

Finally, beginning in 2017-18, quality designations—and consequently, CLASS scores—became linked to financial incentives, including child care tax credits and child care subsidy amounts. The goal was to encourage ECE providers to improve their CLASS score and seek higher quality designations, as predicted under the QRIS model. Conversely, financial consequences were also introduced: centers deemed "unsatisfactory" for two out of three years would lose their public funding.

The Landscape of Publicly-Funded ECE in Louisiana

The LDE classifies center-based ECE providers into the following categories: *Type I*, or programs owned or operated by a church or religious organization that do not take any public funding; *Type II*, or programs that either take no public funding or only take food and nutrition funding; *Type III*, or programs that take public funding (CCAP, NSECD, Head Start/Early Head Start, etc.); and *School-Based*, which operate in schools and offer before and after school care.

Because Louisiana's QRIS requires participation from all providers receiving public dollars, center-based settings must be classified as Type III if they wish to benefit from the financial incentives associated with quality ratings. Conversely, Type I and Type II programs do not receive CLASS scores, though they are also not eligible for the state's financial incentives and must still comply with state regulation. Because this report aims to understand levers for quality improvement as measured by QRIS, programs that do not participate in QRIS (i.e. Type I and Type II) are not considered here.

Publicly-funded programs comprise school-based programs (44.1%), child care centers (42.8%), and Head Start programs (13.2%). Figure 3, at right, depicts the distribution of

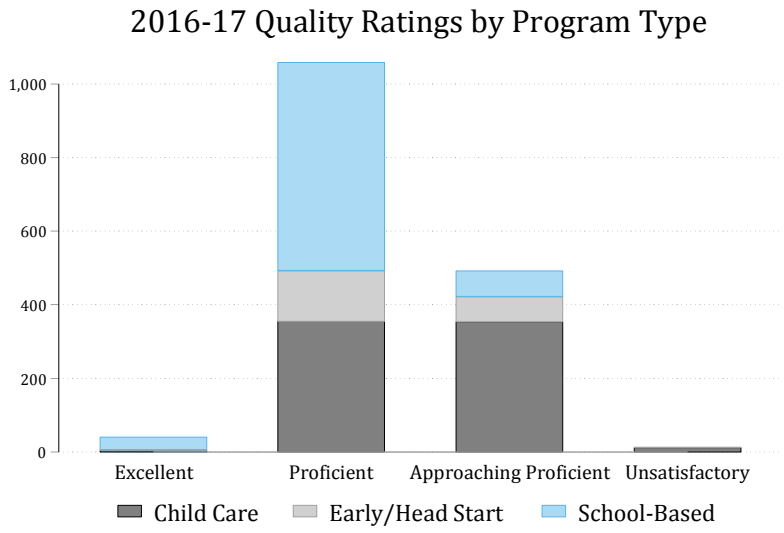

Figure 3. Publicly-funded ECE programs (Type III and school-based) in Louisiana during the 2016-17 school year, by quality rating and program type. The average CLASS score underlying these ratings was 4.86.

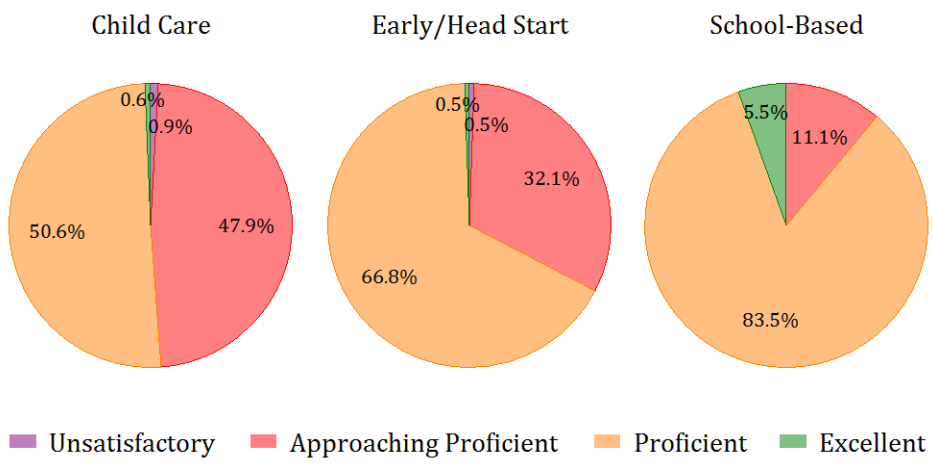

Figure 4. Quality rating designations in 2016-17, by program type. Ratings are listed in legend in ascending order, where "Unsatisfactory" represents the lowest quality designation, and "Excellent" denotes the highest designation. Average CLASS scores across program type were 4.50, 4.72, and 5.22, respectively (out of 7).

programs both by program type, and by quality rating. Encouragingly, two-thirds of programs overall were in the top two quality designations, although nearly a third of programs received quality ratings that were less desirable. However, child care centers and Head Start programs were disproportionately represented in lower-quality ratings, while school-based preschools composed nearly the entirety of "Excellent" programs.

Figure 4 (above) depicts these disparities more clearly: nearly half of child care centers and nearly a third of Head Start programs were either "Approaching Proficient" or "Unsatisfactory," relative to just 11.1% of school-based programs. In contrast, while 5.5% of school-based programs were rated "Excellent," not even one percent of Head Start or child care centers received this quality rating.

It is not evident what might explain these sector differences in quality, although the differences reflected here generally align with descriptive trends at the national level (Bassok, Fitzpatrick, Greenberg, & Loeb, 2016). One hypothesis is that classrooms in school-based programs may be required to also comply with K-12 education standards for educational quality. As described in Table 1 below, teachers in school-based programs appeared far more likely to hold a bachelor's or master's degree. In contrast, the majority of teachers in Louisiana child care centers did not have a

Table 1. Louisiana's ECE workforce by program type

	Child Care	Early/Head Start	School-Based
% of Teachers with…			
Master's degree	2.45	4.99	21.53
Bachelor's degree	10.18	47.34	69.99
Associate degree	8.38	31.39	0.76
CDA	23.53	12.65	0.8
No Degree	55.44	3.63	7.64
LA certification	25.28	19.75	21.46

Notes. CDA = Child Development Associate, a professional certificate. The first five categories are mutually exclusive; certification can overlap with other education levels, and can include Louisiana's Early Childhood Ancillary Certificate.

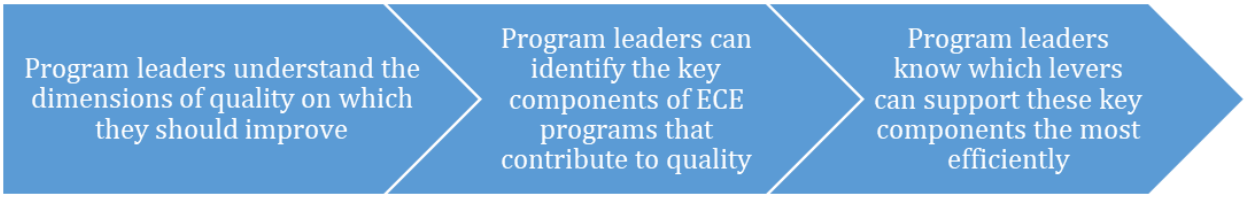

Figure 5. Theory of action, and underlying assumptions, for how program leaders engage in quality improvement decisions.

college degree. However, there may also be other systematic differences, such as the resources available to different programs, that preclude different types of programs from quality improvement.

These disparities call into question the strategies that program leaders, particularly in non-school-based settings, are able to employ in response to state-defined goals for quality improvement. Figure 5, above, provides a simplified model for thinking about what information program leaders need in order to make effective quality improvement decisions. Specifically, quality improvement assumes that (1) program leaders understand the dimensions of quality on which they should improve, (2) program leaders understand the malleable components of ECE programs that should be targeted to foster quality improvement, and (3) program leaders are aware of the available classroom supports that target these components.

Knowledge of these three pieces can help programs to efficiently invest in supports that best foster quality improvement. In the Louisiana context, the first component of this model is satisfied: Under QRIS, programs know which domains of CLASS require improvement to obtain a higher quality rating. The following sections discuss information asymmetries in the second and third components.

How Can Policies Influence Quality?

Inherently, systems such as QRIS aim to promote the quality of classroom interactions. Indeed, researchers have credited the success of pioneering preschool programs (Campbell et al., 2012; Schweinhart et al., 2005) in part to the richness of the interactions and processes that participating children received in these settings (Phillips et al., 2017). However, with Louisiana's QRIS as a notable exception, most state policies do not target classroom processes directly. Instead, they target various components of the ECE program that are hypothesized to support quality classroom processes, such as teachers or classroom resources. Connors (2016) provides a theoretical model, described below and in Figure 6, for how effective policy influences not only children's learning environment but also teachers' professional environment.

Classroom structural quality. Classroom structural quality can be described as "the quality of [classroom] resources … and the organization of those resources" (Connors, 2016, p. 34). Structural quality features have been traditionally targeted by policy, given that they are easy to assess and regulate. For example, licensing regulations monitor features such as developmentally appropriate materials, teachers' formal education, and teacher-child ratios (Payne, 2011).

Classroom structural quality is conceived as the "foundation" for the supportive and stimulating processes that characterize a high-quality ECE classroom (Connors, 2016; Phillips, McCartney, & Sussman, 2008). In other words, these features establish the context in which warm and engaging teacher-child interactions occur. Unfortunately, because structural features indirectly affect

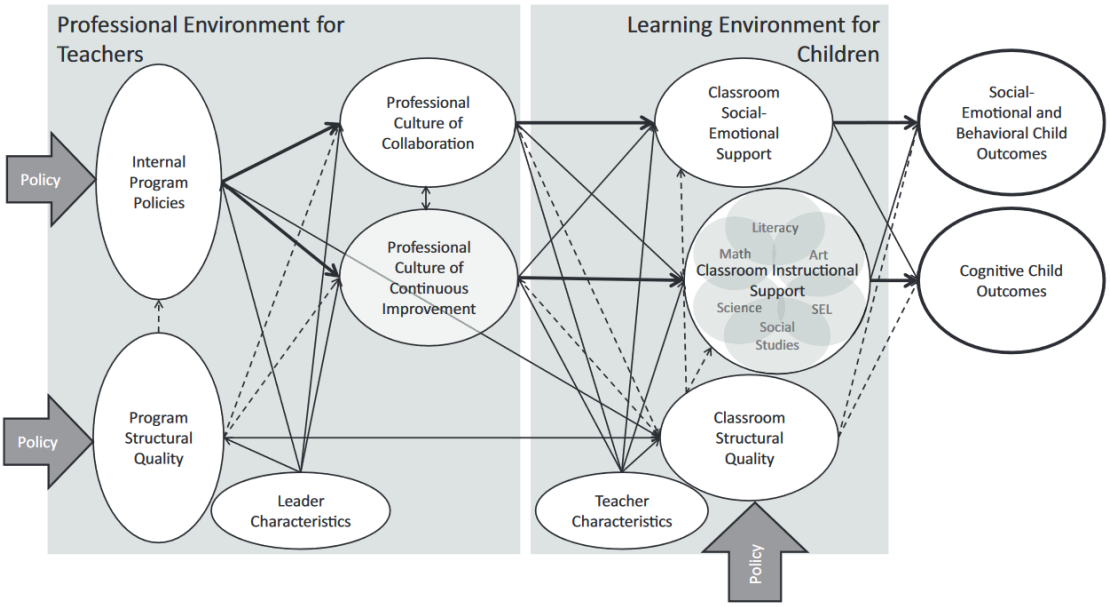

Figure 6. Theoretical model of how policy affects classroom quality. Policy improves classroom quality and child outcomes through its influence on teachers' professional environment and children's learning environment. Arrows are weighted (e.g., dashed, bolded) according to the strength of their hypothesized relationship. Adapted from Connors (2016).

children, *changes* in structural features are not very predictive of positive changes in child outcomes (Lin & Magnuson, 2018; Pianta et al., 2005), or even changes in classroom process quality (Bassok & Galdo, 2016; Sabol & Pianta, 2014). Thus, policies targeting structural features are a necessary but not sufficient condition for supporting rich and engaging classroom processes, and policies aiming to encourage *improvement* in quality should target characteristics that better predict positive change in classroom process quality.

Internal program policies. Internal program policies, sometimes referred to as professional culture, "reflect formalized processes among teachers, administrators and staff … [that] are intentionally designed to help teachers create and improve high-quality learning environments for children" (Connors, 2016, p. 37). Professional culture describes the processes that occur in teachers' professional environment, much in the same way classroom process quality captures processes in children's learning environment. Examples of policies that influence professional culture include holding regular staff meetings, supporting teachers through training and coaching, and providing paid planning time to teachers (Connors, 2016).

As described in Figure 6, there are two components of professional culture: a *culture of collaboration*, which "describes the norms … governing the relationships that exist between and among teachers, staff and administrators" (Connors, 2016, p. 38); and a *culture of continuous improvement*, which focuses on "norms regarding adults' reflection on their practice, engagement in ongoing learning, experimentation with new strategies and ideas, and commitment to striving higher" (Connors, 2016, p. 38). Together, programs with high-quality professional cultures are teaching environments where teachers are engaged with learning and implementing teaching strategies, and where teachers feel supported in their teaching practice.

Program structural quality. Program structural quality encompasses "the people, resources, materials, physical environment, and organization of time that directly support teachers" (Connors,

2016, p. 36). Examples of policies that affect program structural quality features include generous staff pay and benefits, and directors' education and experience.

Analogous to classroom structural quality, *program* structural quality is conceived as foundational for processes that stimulate teachers' learning in their professional environment. For example, a teacher may acknowledge the importance of continuously refining her teaching practice, but may be unable to engage in these opportunities if available program resources and supports are inadequate. Expanding the definition of ECE quality to incorporate not only the child's learning environment but also the teacher's professional environment acknowledges their interrelated and parallel relationship—specifically, that features of one environment can be leveraged to support the other (Knapp, 2003).

Program Responses to Quality Improvement

Recall that QRIS provide program leaders with information on what areas of quality they should improve, but do not prescribe specific actions for *how* to improve. If program leaders lack information on how and which program components contribute to quality, then they may face challenges in making decisions that effectively foster quality improvement. To address this information asymmetry, the LDE provides general guidance on three types of investments: (1) classroom resources, (2) teacher preparation and certification, and (3) ongoing professional development. These types of investments, described below, are directly related to the three policy arrows in Connors' theoretical model (2016), and they can be adopted both by individual programs and across community networks.

Classroom resources. Classroom resources, such as curricula, materials, and manipulatives, are often considered the vehicle for selecting and implementing opportunities and activities for meaningful interactions with students (Ansari & Winsler, 2014; Domitrovich et al., 2009). However, only recently have researchers begun to consider what curricula content constitutes "high-quality" (Clements, 2007). Further, the effectiveness of "research-based" curricula varies, and even well-designed curricula that have the most potential to effectively improve quality are not typically used in ECE (Duncan et al., 2015). Nevertheless, there is consensus that ECE curricula should be age-appropriate and promote early skill development (Ansari & Winsler, 2014; Duncan et al., 2015).

In Louisiana, ECE program leaders can choose any curricula they deem best for the children and families they serve. That said, the LDE encourages curricula choices that are appropriate for young children's development and are grounded in helping teachers provide quality interactions and instruction. The LDE evaluates curricula on a number of standards, including *cultural and linguistic sensitivity*, *options for family engagement*, *ease of use by staff*, *quality and complexity of learning activities*, *embedded assessment*, *potential for individualizing instruction*, and *developmental appropriateness*. Curricula meeting these standards are classified as "Tier 1" by the LDE. At present, the LDE has approved Tier 1 ECE curricula from 12 different publishers.

To encourage use of Tier 1 curricula among publicly-funded programs, the LDE offers reimbursement of 80% of costs toward purchase of Tier 1 curricular materials (maximum two curriculum kits totaling $4,000). It also maintains a vendor listing of different curricula from which program leaders can choose. Curricular packages are categorized by tier, and information about instruction, coaching, implementation, and estimated costs are provided in this listing. However, most curricula packages do not have quantifiable information on the projected benefits of curricular use, particularly with respect to the CLASS. Beyond its sponsoring of Tier 1 curricula, the LDE provides no additional guidance for selecting specific curricular packages.

Teacher preparation. The content knowledge and skills that high-quality teachers bring to the classroom have been widely recognized as an important contributor to young children's learning

(Hamre, 2014; Hamre et al., 2014; Hatfield, Hestenes, Kintner-Duffy, & O'Brien, 2013). There are three general approaches to the education and training of ECE professionals: *formal education*, which is outside the scope of this report[1]; *in-service training*, which is the focus of the section below; and *credentialing*, which is the focus of this section.

The purpose of credentials is to signal that a teacher has met a set of competencies related to nurturing the emotional, physical, intellectual, and social development of children; and that the teacher has knowledge for putting these standards into practice. The most common ECE credential is the Child Development Associate (CDA), though there are others recognized both nationally and locally (Maxwell, Field, & Clifford, 2006).

Louisiana is becoming increasingly responsive to setting expectations and requirements for teacher preparation and certification. Beginning 2019, Louisiana intends to require teachers in publicly-funded ECE to hold an Early Childhood Ancillary Certificate, which is a credential specific to the state of Louisiana. To obtain this certificate, teachers would have to hold either a CDA in Infant/Toddler or Preschool, or a technical diploma or certificate in an early childhood related field from an accredited technical or community college. The LDE would also require this certificate to be conferred by a pre-approved Louisiana training program or institution. Finally, the LDE provides programs with financial support to help their teachers obtain Early Childhood Ancillary Certificates, through competitive *Louisiana Pathways Scholarships*, which are funded by the federal Child Care and Development Block Grant (CCDBG).

Ongoing professional development. Ongoing professional development (PD) opportunities are the primary pathway through which program leaders can invest in professional culture. PD supports aim to provide teachers with feedback that helps them to improve their skills and teaching strategies, that they can then apply in their classroom (Sheridan, Edwards, Marvin, & Knoche, 2009). However, there is wide variation in the design, content, and delivery of different PD programs (Schachter, 2015). PD can include in-service training, coaching, and connecting teachers with other professionals to form a learning community. In contrast to teacher preparation, which focus on the foundational skills and knowledge teachers bring to the classroom, PD supports teachers by targeting programs' culture of collaboration and continuous improvement.

Meta-analyses generally conclude that PD in ECE can effectively change teachers' practices in ways that lead to improved student outcomes (Fukkink & Lont, 2007; Kraft, Blazar, & Hogan, 2017). One study in particular (Early, Maxwell, Ponder, & Pan, 2017) demonstrates that PD can have small effects on subsequent CLASS observations. For PD to be effective, it must be aligned, ongoing, and targeted (Hamre, Partee, & Mulcahy, 2017).

Recognizing the importance of these PD components, the LDE has explicitly expressed its expectation for ECE teachers to participate in PD. The LDE supports PD opportunities in several ways. First, the LDE hosts an annual *Teacher Leader Summit* offering various professional development opportunities. Second, both at the summit and throughout the school year, the LDE recommends networking with other ECE and early elementary professionals, through which ideas, knowledge, and new practices can be exchanged. Last, the LDE provides information on its website

[1] Although formal education—which encompasses both pre-service training (usually, teacher candidates preparing to enter the workforce) and post-service training (coursework taken while employed), and focuses on foundational knowledge important for teaching young children—are an important component of teacher preparation, there is no program-level policy that would adequately improve program quality, without structural changes to higher education preparation programs or the ECE teacher labor market at large.

as guidance to parishes, such as how to contact local child care resource and referral agencies, and links to free, online courses for training.

Beyond these freely available supports, the LDE encourages program leaders to review their CLASS results and consider PD options from a vendor catalog, depending on what "best meet the needs of your teachers.[2]" Although the catalog contains rich information on professional development options, including cost estimates and whether they are aligned with state-approved curricula and offer ongoing coaching, no additional guidance on which professional development programs to purchase is provided.

Putting It All Together

Figure 7 summarizes how Louisiana's policy context and the existing literature inform my theory of action for how program leaders engage in quality improvement decisions. Louisiana has undertaken unprecedented efforts for providing clarity on the important dimensions of quality classroom processes, the components of ECE programs that are amenable to change, as well as an extensive list of all classroom supports available to program leaders. Yet one challenge that remains unaddressed is an understanding of how programs can invest both effectively and efficiently in

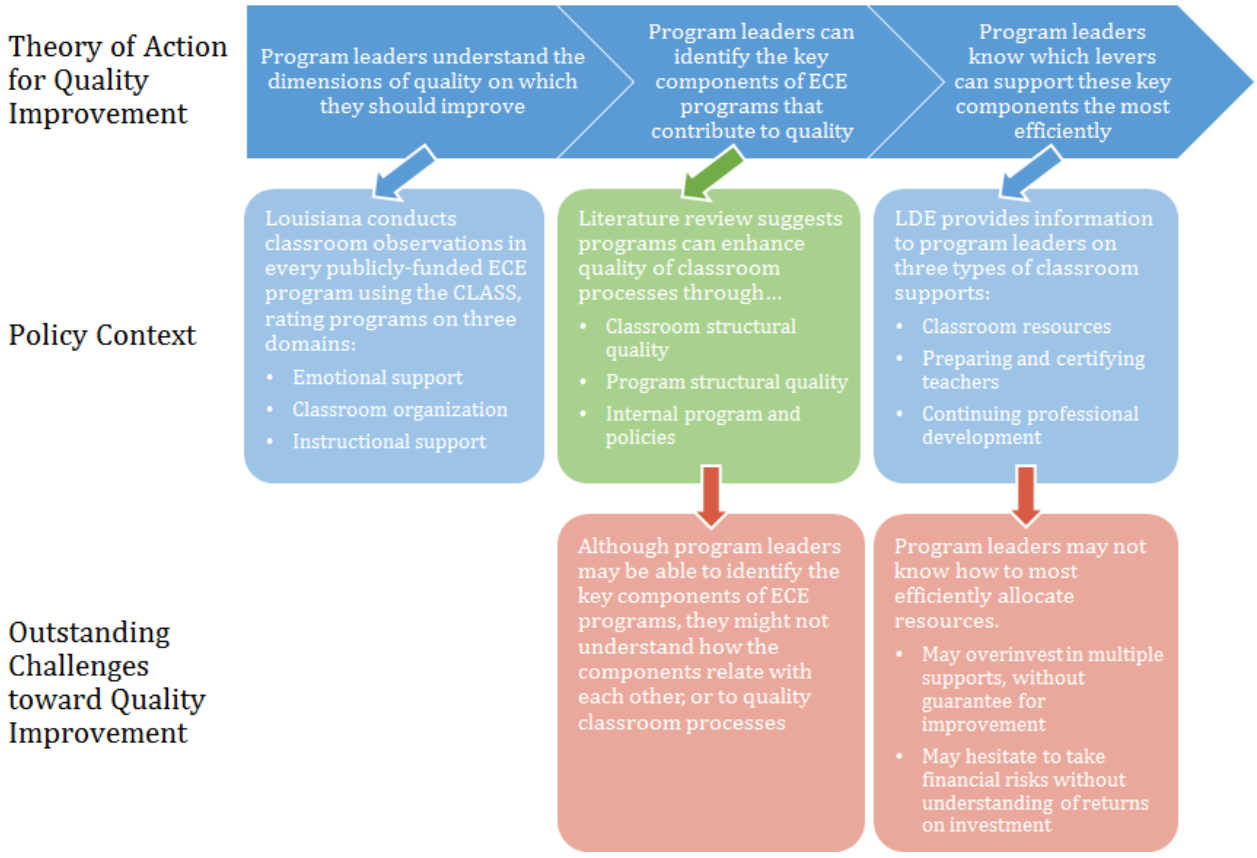

Figure 7. Summary of literature: What challenges do Louisiana ECE programs face toward quality improvement?

[2] Retrieved 1 April, 2018 from https://www.louisianabelieves.com/docs/default-source/early-childhood/guide-to-early-childhood-curriculum-assessments-and-professional-development.pdf

classroom supports. Indeed, two outcomes may occur: Either program leaders may employ a kitchen-sink approach and overinvest in classroom supports, such that the same improvement to quality could have been achieved through a more strategic choice of investments; or they may hesitate to make improvements and take financial risks at all, without a better understanding of the returns to investment they should expect.

Thus, while information on *choices* for classroom supports is undoubtedly helpful, more can be learned about the relative effectiveness of various classroom supports, particularly in relation to their costs. As such, the first part of this report is concerned with understanding the quality improvement efforts that ECE program leaders employ, as well as identifying which of these classroom supports can effectively foster quality improvement. Based on this analysis, the second half of the report evaluates how *cost*-effective these levers are, as well as whether effective classroom supports work especially well for programs in underserved communities, thereby promoting equitable outcomes.

Finally, although quality improvement decisions are enacted locally in response to state-defined goals, the LDE can still benefit from understanding *who* engages in effective quality improvement decisions, as it can help them better target resources to increase access to improvement levers, particularly if the state is concerned that programs in communities with fewer resources are being forced to make tradeoffs between program costs and quality (Bassok & Galdo, 2016; Hatfield, Lower, Cassidy, & Faldowski, 2015; Hotz & Xiao, 2011). For example, while costly investments may generate developmental benefits for children, they may simultaneously increase a program's operating costs and potentially increase the price for care. Identifying effective classroom supports for fostering quality improvement, and ensuring program leaders can equitably access these levers without bearing the financial burden of improvement, helps the LDE promote its mission of quality ECE for all children.

PART I: IDENTIFYING EFFECTIVE QUALITY IMPROVEMENT LEVERS

Methodology

Data

I leverage data from several sources. First, I analyze survey responses from the *Louisiana 2017 Child Care Market Rate Survey* (MRS), which contains rich data on the revenues and costs incurred by ECE programs during the 2016 calendar year (see Appendix B for items used in this analysis). These data allow me to identify both the classroom supports programs chose to invest in, and total expenditures toward each of these improvement efforts. I then merge in a two-year panel of *administrative data* from the LDE, which contain scores from the state's systematic classroom observations, as well as programs' QRIS designation. This allows me to assess the extent to which programs' investments in classroom supports translated into quality improvement. Finally, I supplement my data with parish-level demographic information from the *2016 American Community Survey 5-year estimates* (ACS). I use these data to explore what programs were able to access classroom supports, and I also include them as covariates when estimating the effectiveness of an improvement lever.

My final analytic sample comprises 482 Type III programs. A summary of data sources is provided below, and a detailed explanation of the final sample selection is provided in Appendix C.

Louisiana Child Care Market Rate Survey. In accordance with federal requirements, the MRS was administered so that the LDE could set adequate reimbursement rates that would allow subsidized families equal access to the full range of child care available to families not receiving child care subsidies. Focusing on Type III programs—those receiving quality ratings—I analyze responses to the *Child Care Center and School-Based Child Care* form of this survey. Respondents were invited throughout the 2017 calendar year to complete the survey online; respondents who did not complete this version were sent email reminders, as well as two mailings of a printed version of the survey. The survey, which was to be completed by a director, assistant director, or principal, asked questions regarding financial decisions from January 1st, 2016 to December 31st, 2016. These data contain rich information on enrollment, prices charged to paying families and revenues earned, and expenditures related to both everyday operations (such as insurance and rental costs) and improvement efforts (such as professional development, curricula, and teachers' wages).

LDE administrative data. Beginning the 2015-16 school year, the LDE conducted systematic observations of all classrooms in publicly-funded ECE programs using the CLASS. These administrative data, which span two years, include the overall CLASS score, CLASS domain scores, QRIS designation, percent composition of teachers by education level and certification, teacher-child ratios, and curriculum quality.

American Community Survey. The ACS is administered annually by the US Census Bureau, and contains a rich set of demographic variables measured at various levels of geography (e.g. zip code, parish). I consider the following variables in my analysis: median household income, percent non-Hispanic Black, percent Hispanic, number of children aged five or under, maternal labor force participation rate, and unemployment rate.

Data strengths and limitations. In general, these data are more exhaustive than most readily accessible datasets, which often do not collect financial information from ECE programs at scale.

However, the primary advantage of these data is that they directly describe Louisiana ECE programs and their quality improvement decisions, such that effectiveness estimates can be derived directly rather than inferred from the literature.

Nevertheless, there are two notable limitations to these data. First, responses from this survey are self-reported, and the quality of responses depends on accuracy of reporting (e.g., a respondent might report an annual wage, instead of an hourly wage), as well as the quality of program leaders' recordkeeping (e.g., respondents might provide a 'best-guess' estimate if they did not have an exact figure at hand). With the exception of outlier data points, these hypotheses are hard to empirically assess.

A second limitation is that survey data lend themselves to response bias. My missing data analyses (see Appendix C) suggests that child care centers were more likely to respond to the survey, relative to Head Start programs. Programs with a greater proportion of teachers with Associate degrees, and programs with a smaller share of teachers without certification or a college degree, also had greater odds of responding to the survey. I address these concerns when computing my effectiveness estimates (see Appendix C).

Measures

Classroom resources. Program leaders listed the specific classroom materials they purchased (see Appendix B, question 14). The primary challenge with using this item was the wide range of responses, given the open-endedness of the question. Thus, I code these responses and construct a dichotomous variable that indicates whether respondents mentioned any of the Tier 1 curricular packages.

In addition, I consider a measure of participation in curricula-related training (see Appendix B, question 16), to proxy for the implementation of curricular materials, above and beyond simply purchasing materials. Simple descriptive statistics suggest the majority of respondents using Tier 1 curricula participated in training (81%), compared to 52% of respondents not using Tier 1 curricula.

Teacher preparation. For each program I observe the percentage of teachers with a given education level (e.g., percent of teachers with a CDA), but not the exact number of teachers employed. My variables of interest are the compositions of teachers with a CDA, and of teachers with certification (which can include the Early Childhood Ancillary Certificate). I construct a dichotomous variable that is equal to one if the percentage of teachers is greater than or equal to 50%, and zero if the percentage of teachers is under 50%.

Professional development. Program leaders indicated whether they participated in various PD opportunities (see Appendix B, question 16). Seven items were listed in this question. The first item related to trainings specific to curricula use, which I include in my evaluation of curricula rather than professional development. The remaining six items describe PD opportunities that, although vary in structure and design, all incorporate the CLASS measure as part of training; the exception is GOLD training, which offers training and coaching for effectively using the state's portfolio-like system for assessing children's competencies, Teaching Strategies GOLD. Table 2 (see next page) describes the differences across the PD programs, as well as how PD measures were constructed.

Table 2. Professional development measures

PD Program	Variable Definition	Description of PD Program
Making the Most of Classroom Interactions (MMCI; Hamre et al., 2012)	1 = Reported staff participated in MMCI; 0 = Reported no staff participated in MMCI	MMCI is a group coaching modeling, where an in-house CLASS coach helps teachers develop awareness and enact change in the classroom. MMCI emphasizes *face-to-face* coaching in a *group setting*.
My Teaching Partner (MTP; Pianta, Mashburn, Downer, Hamre, & Justice, 2008)	1 = Reported staff participated in MTP; 0 = Reported no staff participated in MTP	MTP is a system of professional development supports that primarily consists of (1) a video library with examples of best practices for classroom interactions; (2) a three-credit college course focused on identifying and applying effective interactions to the classroom; and (3) *web-mediated, 1-on-1 video coaching*.
MyTeachStone	1 = Reported staff used MyTeachStone; 0 = Reported no staff were using MyTeachStone	MyTeachStone is an online platform for measuring, mentoring, and sustaining a culture of teaching excellence. It is recommended for programs seeking to streamline CLASS implementation and facilitate use of assessment data to inform practice. It is also a platform for online coaching and fosters an online professional learning community to learn from peers.
CLASS PreK observer reliability training (CLASS PreK training)	1 = Reported staff participated in CLASS PreK training; 0 = Reported no staff participated in CLASS PreK training	CLASS observation training is designed to help organizations build capacity to collect and use CLASS teaching assessment data. CLASS training helps teachers (and coaches) develop CLASS fluency, helping them to identify and apply best practices in classrooms.
CLASS Toddler observer reliability training (CLASS Toddler training)	1 = Reported staff participated in CLASS Toddler training; 0 = Reported no staff participated in CLASS Toddler training	CLASS observation training is designed to help organizations build capacity to collect and use CLASS teaching assessment data. CLASS training helps teachers (and coaches) develop CLASS fluency, helping them to identify and apply best practices in classrooms.
GOLD Training	1 = Reported staff participated in GOLD training; 0 = Reported no staff participated in GOLD training	GOLD is an ongoing, observation-based assessment that supports effective teaching and assessment. It is an online tool that creates developmental profile of each child, and can generate comprehensive reports for family members and stakeholders. GOLD trainings help teachers with implementation and planning, as well as increase teachers' data literacy.

Empirical Strategy

This section describes how I identify usage of classroom supports, and how I evaluate the effectiveness of these levers.

Determination of access. To understand selection into use of various classroom supports, which I use as my description of the status quo, I fit the following logistic regression:

$$(Policy = 1)_{ipt} = \alpha + \sum_{x \in \bar{\boldsymbol{X}}_i} \beta_3 x + \sum_{p \in \bar{\boldsymbol{P}}_p} \beta_4 p + \epsilon_{ipt} \quad (1)$$

where $\bar{\boldsymbol{X}}_i$ is a series of program-level characteristics (e.g. Head Start, % of teachers with at least a bachelor's degree); and $\bar{\boldsymbol{P}}_p$ is a series of parish-level characteristics; standard errors are clustered at the parish level. Covariates are input simultaneously to adjust for potential interactive effects between different program- and parish-level characteristics. The slope coefficients in this model determine whether, under the status quo, some types of programs (e.g., those in higher-income parishes) had greater odds of adopting a classroom support.

Estimation of effectiveness. To provide some intuition for identifying the expected effectiveness of a classroom support, consider two programs receiving comparable CLASS scores in the 2015-16 school year. Based on this score, each program leader is faced with decisions on how to marshal resources toward quality improvement, which is then reflected in their 2016-17 CLASS score. Thus, part of the 2016-17 score is determined by programs' previous score, and the estimate we are interested in is: above and beyond these determinants, in what ways is the decision to invest in a given classroom support associated with higher quality outcomes?

However, research has demonstrated that program characteristics, such as program type, are related to program quality (Bassok et al., 2016), and these may also be related to decisions to use or not use classroom supports. In addition, community characteristics—above and beyond programs themselves—are associated with quality (Bassok & Galdo, 2016; Hatfield et al., 2015) and may influence the kinds of decisions programs tend to make. This is particularly likely under Louisiana's model of early childhood community networks, as programs in the same parish may have access to similar resources.

Following this rationale, the primary regression specification I fit is:

$$CLASS_{ip(t+1)} = \alpha + \beta_1 (Policy = 1)_{ipt} + \beta_2 CLASS_{ipt} + \sum_{x \in \bar{\boldsymbol{X}}_i} \beta_3 x + \sum_{p \in \bar{\boldsymbol{P}}_p} \beta_4 p + \epsilon_{ipt} \quad (2)$$

where $(Policy = 1)_{ipt}$ is a program-level measure for whether a program *i* adopted the classroom support; $CLASS_{ipt}$ and $CLASS_{ip(t+1)}$ represent the 2015-16 and 2016-17 CLASS scores for program *i* in parish *p*, respectively; and $\bar{\boldsymbol{X}}_i$ and $\bar{\boldsymbol{P}}_p$ maintain the same interpretation from Equation (1).[3]

The parameter β_1 is my estimate of effectiveness, and represents the expected difference in average CLASS score in 2016-17 between those that invested in the classroom support $(Policy = 1)_{ipt}$ and those that did not $(Policy = 0)_{ipt}$, adjusting for what CLASS scores programs received in the prior year, as well as program and parish characteristics. Specification checks, which I present in

[3] Two additional notes. First, in all models, standard errors are clustered at the parish level. This decision is particularly important here, given that Louisiana has established early childhood community networks as part of the provisions of Act 3, so we may expect correlated errors among programs within the same parish. Second, estimates also take into consideration missing data. This approach is described in Appendix C.

Appendix D,[4] confirm that estimates are particularly sensitive to parish characteristics, suggesting the need to include these variables in estimation.

A final note is that although my specification considers a number of potential confounding variables that might influence the relationship between use of classroom supports and quality outcomes, my effectiveness estimates are associations at best and should not be interpreted as causal effects. Rather, the results presented below should be interpreted as program-level practices that *correlate* with improvement.

Results

Access

Classroom resources and curricula. As described in Figure 8 (at right), 70% of repsondents mentioned a state-approved, Tier 1 curricula. Head Start programs were more likely to report purchase of Tier 1 curricula (91%) relative to child care centers (69%). Five percent of respondents listed a non-Tier 1 curriculum, and 25% did not list a curricula package by name. The most frequently mentioned Tier 1 curricula were Teaching Strategies' *The Creative Curriculum* and Frog Street Press' *Frog Street* curricula.

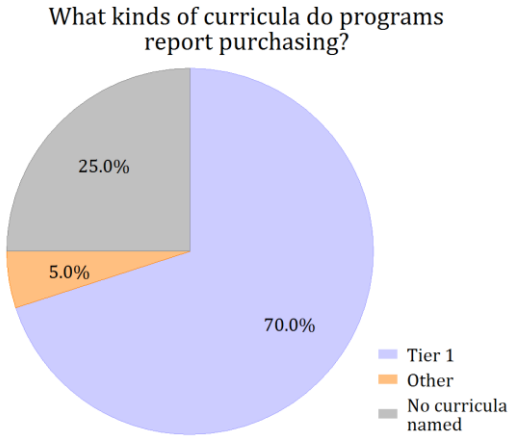

Figure 8. What kinds of curricula do programs report purchasing? Based on N=200 respondents.

On the whole, there were few significant predictors of selection into curricula use (see Appendix D). Programs in communities serving more Black non-Hispanic families had fewer odds for adopting Tier 1 curricula, relative to programs in communities serving fewer Black non-Hispanic families.

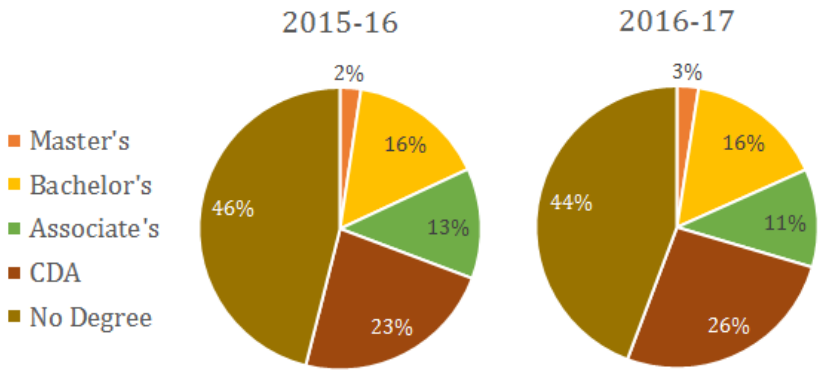

Figure 9. What does the composition of teachers look like in the average ECE program? Based on N=426 programs. No Degree = High school diploma or equivalent.

Teacher preparation and certification. Figure 9 at left describes the composition of teachers by education level, in both the 2015-16 and 2016-17 school years. There appeared to be small shifts toward teachers with the Child Development Associate (CDA) credential. At the same time, the proportion of teachers without a college degree or certificate and the proportion of teachers with an Associate degree decreased very slightly.

[4] In Appendix D, I discuss a fixed effects approach for addressing additional *unobserved* confounding variables. Therein, I justify my rationale for preferring Equation (2).

What PD opportunities do programs participate in?

30%	*participated in* ***Making the Most of Classroom Interactions***
15%	*participated in* ***My Teaching Partner***
72%	*participated in* ***MyTeachStone***
75%	*participated in* ***CLASS PreK observer reliability training***
69%	*participated in* ***CLASS Toddler observer reliability training***
94%	*participated in* ***GOLD training***

Figure 10. What kinds of professional development do programs report participating in? Number of respondents ranged from N=372 to N=395 across survey questions.

Encouragingly, the proportion of teachers certified to teach in Louisiana increased during this period as well. In the 2015-16 school year, the average program had 25% of its teachers certified to teach, while in the 2016-17 year the average program had 31% of its teachers with Louisiana teaching certification.

As summarized in Appendix D, there were more predictors of programs having 50% or more of its workforce composition credentialed. Higher baseline CLASS scores and serving communities with higher maternal labor force participation rates were associated with greater odds of having more than 50% of teachers credentialed.

Professional development. Among all the PD interventions examined in this report, *GOLD Training* had the highest rate of participation (see Figure 10 above). Centers with a smaller share of teachers with a bachelor's degree had greater odds of participating in GOLD training (see Appendix D). *CLASS PreK training* and *MyTeachStone* were also widely adopted, and both higher quality programs and programs in higher income areas had greater odds of participating in these PD programs. MTP appeared to be the least frequently selected PD; however, with the exception of parish unemployment rates, there were no variables that systematically predicted use of MTP. Approximately 7.3% of respondents reported participating in only one of these programs; 6.2% selected all six choices. While correlations between many of the PD programs ranged between .1 and .3, the most highly correlated PD options were the two CLASS reliability trainings (r = .44), and MMCI and MTP (r = .30).

Effectiveness

Classroom resources and curricula. Programs that reported purchasing a Tier 1 curricula seemed to have higher CLASS scores in 2016-17 than those that did not purchase Tier 1 curricula (see Figure 11 on next page). However, this estimate was not statistically significant, even at the 10% level. Participation in training related to curricula also appeared to not be related to higher CLASS scores in the 2016-17 year.

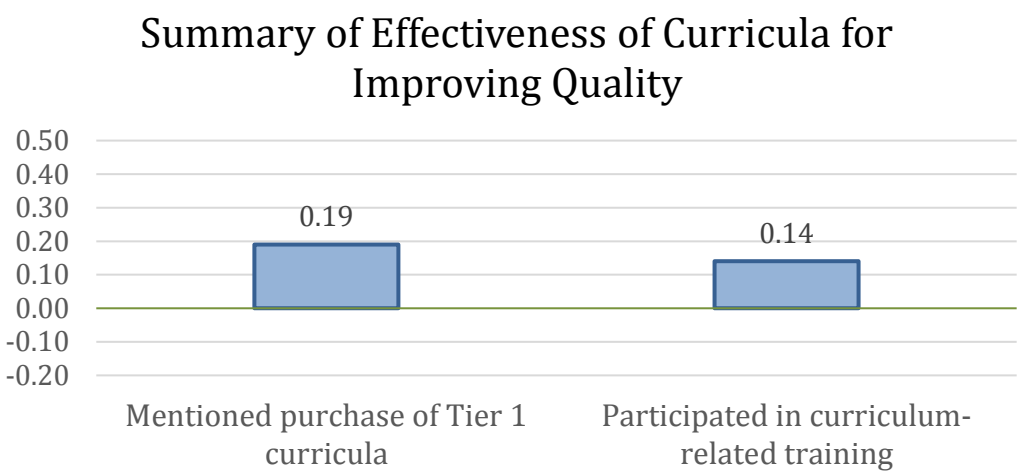

Figure 11. Summary of effectiveness of curricula for improving quality. Each bar represents the coefficient from the same regression, where CLASS 2016-17 scores are regressed on a dichotomous indicator for each curricula variable. Results are adjusted for 2015-16 CLASS score, program type, % of teachers with a bachelor's degree or higher, missingness in curricula variable, and above/below state median on unemployment and maternal labor force participation rates, number of children five and under, percent non-Hispanic black, percent Hispanic, and median household income variables. Regression based on N=482 programs; see Appendix D, Table D3 for regression results across all models, for each curricula variable. Significance: † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. (No results displayed in this graph were statistically significant.)

Although classroom resources are an important component of program leaders' choice set for quality improvement, it is also true that curricula is conceived as a classroom structural quality feature (Connors, 2016), so this lack of findings is perhaps less surprising. However, these results should not be interpreted as dismissing the potential for curricula for two reasons. First, my measures account only for *purchase* of materials, and not use or implementation of Tier 1 curricula. Second, the only outcome considered in this analysis is CLASS scores. It may be, for example, that Tier 1 curricula have stronger effects on other outcomes, such as literacy or math. Nevertheless, because CLASS is a policy-relevant measure of quality for the LDE, the lack of effectiveness observed in the present context is sufficient justification to exclude curricula as a potential policy option moving forward.

Teacher preparation and certification. Figure 12 (on next page) suggests that having more teachers certified to teach in Louisiana seemed slightly related to higher CLASS scores; however, this effect was small in magnitude and not significant. This result is difficult to interpret: Although the state's Early Childhood Ancillary Certificate is included in this category, it also includes other types of teaching certification. As such, because the CDA is the primary pathway toward receiving the Early Childhood Ancillary Certificate, I focus my interpretations on the CDA result below.

Programs with 50% or more of its teachers with the professional CDA certificate had higher CLASS scores than those with less than half of its teachers: a difference of 0.16 points. This result is encouraging and suggests that credentialing may have a positive effect on CLASS, even after adjusting for program and parish characteristics.

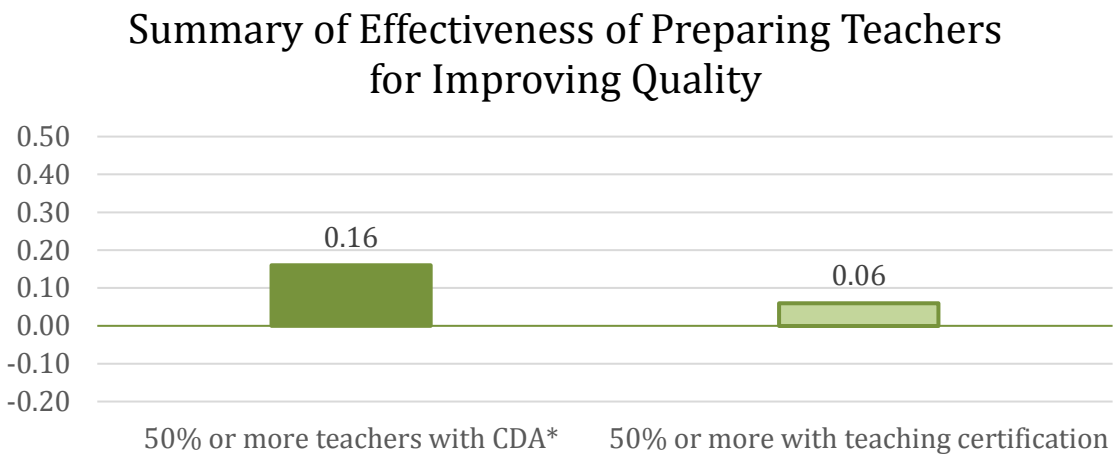

Figure 12. Summary of effectiveness of preparing teachers for improving quality. Each bar represents the coefficient from the same regression, where CLASS 2016-17 scores are regressed on a dichotomous indicator for each teacher preparation variable. Results are adjusted for 2015-16 CLASS score, program type, % of teachers with a bachelor's degree or higher, missingness in teacher preparation variable, and above/below state median on unemployment and maternal labor force participation rates, number of children five and under, percent non-Hispanic black, percent Hispanic, and median household income variables. Regression based on N=482 programs; see Appendix D, Table D4 for regression results across all models, for each teacher preparation variable. Significance: † p < 0.1; * p < 0.05; ** p < 0.01; *** p < 0.001. (Also shaded in dark.)

Professional development. Estimating the effectiveness of a single PD program is tricky since providers could choose to participate in multiple PD programs. Thus, I first regress CLASS scores on all six PD variables simultaneously and estimate the main effects of each PD program. As described in Appendix D, only MMCI, MTP, and CLASS PreK observer training have significant main effects. Second, I examine the sensitivity of these estimates to a model where I include only these three PD variables; I find that these estimates are robust to the exclusion of non-significant PD programs. Finally, I include interactions to examine whether there are synergistic effects for adopting multiple PD programs. I discuss only the interactive model here, but full results are presented in Appendix D.

My findings suggest that participation in MMCI was positively associated with CLASS scores: programs that participated in *only* MMCI (and not MTP or CLASS PreK training) had CLASS scores that were 0.44 points higher, on average, compared to programs that did not participate in any of these PD programs. Note that this effect is the largest observed out of all the classroom support variables studied here. CLASS PreK observer training also had a significant main effect: programs with staff participating only in observer training scored, on average, 0.24 points higher on CLASS than programs not participating in any PD. MTP did not appear to have a significant effect on CLASS scores[5].

Neither the two-way nor three-way interaction terms were significant in this regression, suggesting that there may be diminishing returns to programs participating in multiple PD programs. These

[5] In fact, as described in Appendix D, MTP had trending *negative* associations with CLASS scores, and this relationship was strongest when parish-level covariates were included. However, this estimate is noisy, and is not significant in the final model I interpret here.

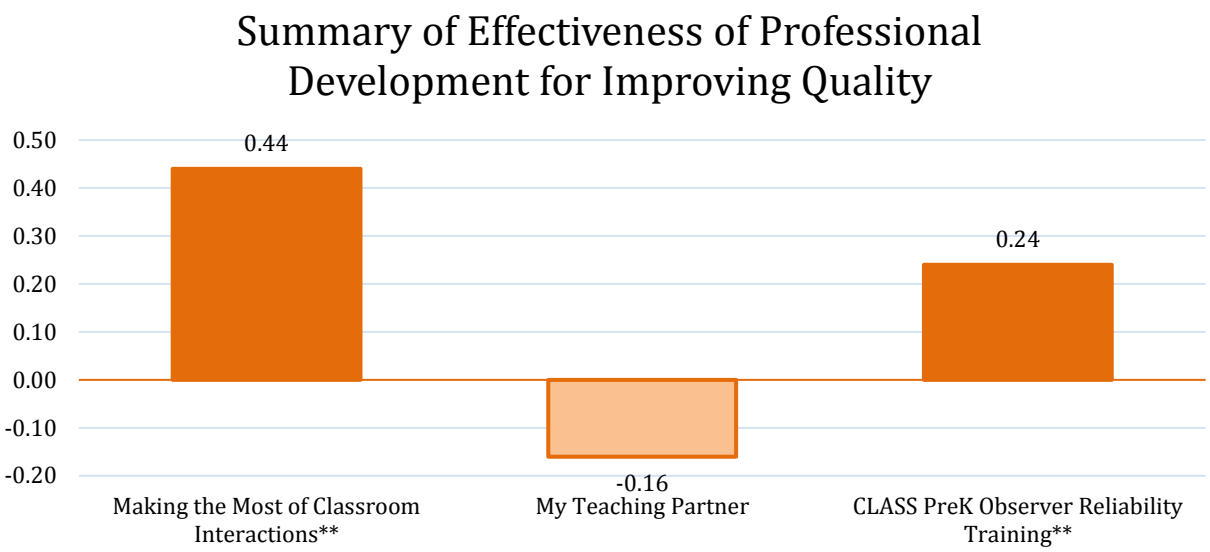

Figure 13. Summary of effectiveness of professional development for improving quality. Each bar represents the coefficient from the same regression, where CLASS 2016-17 scores are regressed on a dichotomous indicator for each PD variable. Results are adjusted for 2015-16 CLASS score, program type, % of teachers with a Bachelor's degree or higher, missingness in PD variable, and above/below state median on unemployment and maternal labor force participation rates, number of children five and under, percent non-Hispanic black, percent Hispanic, and median household income variables. Regression based on N=482 programs; see Appendix D, Table D5 for regression results across all models, for each PD variable. Significance: † p < 0.1; * p < 0.05; ** p < 0.01; *** p < 0.001. (Also shaded in dark.)

patterns make sense from a logistical feasibility perspective, given that multiple PD efforts might consume program resources and overburden teachers.

Summary

My analysis suggests that investments in teachers and supporting their professional environment demonstrate the most potential for fostering quality improvement. However, not all PD are created equal: MMCI, which is a group-coaching model where coaches are situationally embedded in schools, was associated with the largest quality improvement. In contrast, many of the PD programs I examined here, as well as curricula use, did not effectively improve CLASS scores.

It is worth mentioning again that, because it is the policy goal for the LDE to improve CLASS observation scores to encourage higher quality ratings, it makes sense that the most effective levers for improvement were those supports that are most directly aligned with the CLASS framework (i.e., CLASS group coaching, CLASS observer reliability training). However, CLASS scores are not necessarily the only important outcome worth considering, so the LDE should not entirely dismiss classroom supports found in this study to have limited effectiveness. Although outside the scope of the present analysis, the LDE should also consider in future efforts whether these classroom supports are effective at fostering gains in other outcomes, such as children's developmental outcomes.

PART II: FOSTERING QUALITY IMPROVEMENT

To understand in what ways the LDE can play a critical role in supporting program's quality improvement efforts, I begin by first reintroducing the effective improvement levers based on Part I. I then discuss my evaluative criteria for identifying a strong policy option, and then assess each policy option against these criteria.

Policy Options

I propose three policy options the LDE should consider that foster quality improvement among Louisiana ECE programs: implement *Making the Most of Classroom Interactions* at scale, encourage teachers to participate in *CLASS PreK observer reliability training*, and ensure the majority of teachers have earned the Child Development Associate credential. All options are assessed on my evaluative criteria relative to the status quo.

Option 1: Status Quo

A typical policy analysis evaluates policy options relative to present trends. This comparison tells policymakers whether their investments meaningfully improve outcomes above and beyond the status quo; note that this is different than programs doing *nothing*. In other words, any policy option considered here should improve CLASS scores above and beyond increases that may have happened anyway, in the absence of significant investment.

Option 2: Implement Making the Most of Classroom Interactions

Making the Most of Classroom Interactions is a group-coaching model that helps teachers align their practice to the CLASS. The model stresses that coaches mentor teachers in-person and within the school context (e.g., having a scheduled time of week where all teachers meet as a group with the coach). Indeed, my analysis suggests that programs that participated in MMCI (and only MMCI) had on average 0.44 CLASS scores higher than those that participated in no PD, even after adjusting for initial CLASS scores, as well as program and parish characteristics.

This policy option proposes that MMCI be adopted at scale. At present, MMCI is typically adopted at the community network level, rather than at the program level (B. Hamre, personal communication, April 24, 2018). Consequently, the considerations herein explore parish-wide adoption of MMCI, rather than individual program leaders' decisions to participate.

Option 3: Encourage teachers to participate in CLASS PreK reliability training

Although CLASS PreK reliability training is typically used as a research training to calibrate observers to the CLASS rating scale, it is also used as a PD tool for non-research purposes, particularly among program leaders and coaches (B. Hamre, personal communication, April 24, 2018). If teachers know the components of quality interactions and can identify them as implemented in the classroom, then they are better equipped to employ these practices themselves. In other words, the training gives teachers a prototype of what quality teaching should look like.

My analysis finds that programs with staff participating in CLASS PreK training (and only this training) had 0.24 CLASS scores higher, on average, than programs that did not have any staff participating in CLASS training. Thus, this policy option proposes that programs employ at least some teachers or staff who participate in CLASS training.

Option 4: Ensure the majority of teachers have earned the Child Development Associate credential

The CDA credential is perhaps the most widely recognized credential for ECE, and signals both to program leaders and to parents a set of competencies for supporting children's development. My analysis finds that programs with at least 50% of its teachers with the CDA credential have on average 0.16 CLASS scores higher than programs with less than 50% of its teachers with a CDA. Thus, this option is concerned with increasing the share of teachers with a CDA to above 50%.

Evaluative Criteria

An ideal policy option should improve programs' quality *efficiently*, promote *equitable* outcomes, and be *politically feasible* to implement. I operationalize and justify these criteria below:

Table 3. Evaluative criteria for proposed policy options

Criterion	Evaluation	Method	Weight
Cost-Effectiveness (Efficiency)	The extent to which a policy option produces a desired outcome given efficient use of dollars	Compute a *Cost-Effectiveness Ratio*: expected change in CLASS score per dollar spent	40%
Equity	The extent to which a policy option promotes greater quality improvements for programs serving traditionally disadvantaged or underserved communities	Examine heterogeneity in effectiveness of a policy option	40%
Political Feasibility	The extent to which a policy option can garner buy-in from currently non-adopting programs	Literature review; personal communications with experts	20%

Cost-Effectiveness

Cost-effectiveness is a measure of how effective a policy option is at producing a desired outcome (here, quality improvement, measured by the CLASS), in relation to the costs associated with the policy option. I consider the ratio of effectiveness to costs as my primary criterion, rather than just a policy's effectiveness, as every dollar saved can be allocated in a way that more efficiently improves classroom quality. Intuitively, a policy option would have a greater cost-effectiveness ratio if it produces large gains in CLASS score per dollar spent. In contrast, a policy alternative would have a smaller cost-effectiveness ratio if it produces little or negative change in CLASS score. Computation of costs are described in the analysis of each policy option, and effectiveness is derived from Part I of this report.

Equity

A policy option is particularly ideal if it promotes equitable outcomes for all children. For example, a policy option would rate highly on this criterion if programs otherwise serving traditionally disadvantaged or underserved communities saw greater quality improvement resulting from the policy option. In contrast, a policy option would rate low on this criterion if it widened existing educational inequities. To assess equity, I assign each policy option a point-value from one to five based on the description of equitable outcomes provided in Table 4.

Table 4. Summary of equitable outcomes

Predictor Variable	What outcome would promote equity?
Program Characteristics	
2015-16 CLASS	Programs with low baseline quality have greater improvements than programs with high baseline quality.
Parish Characteristics	
Percent Black, non-Hispanic	Programs in parishes serving more racial and ethnic-minority families stand the most to improve
Percent Hispanic	Programs in parishes serving more racial and ethnic-minority families stand the most to improve
Median Household Income	Programs serving families in poverty stand the most to improve

Notes. Equity is determined by splitting the sample based on demographic characteristics (typically, above/below median), and fitting regression in each to compare effectiveness estimates. The full results are presented in Appendix E. Note that estimations include other covariates not listed here; however, only the covariates listed above contribute to evaluations of equity.

Political Feasibility

Political feasibility from the perspective of the LDE is likely to be high for all policy options considered. Instead, I consider feasibility from the perspective of currently non-adopting programs—for example, if they were encouraged by the state or their community network to adopt the policy option. A politically feasible option would easily garner buy in, while a politically infeasible option might face resistance from programs unwilling to change their practice. I assess feasibility on a five-point scale, determined by personal communications with Bridget Hamre, a Research Associate Professor at the University of Virginia involved with the development of both CLASS and related PD programs.

Analysis of Policy Options

Implement Making the Most of Classroom Interactions

Cost-effectiveness. The primary accounting costs associated with MMCI is the training of MMCI coaches. Teachstone—which delivers training for all CLASS PD programs, including MMCI, MTP, and CLASS training—estimates MMCI coach training to be approximately $4000 per coach (K. Kehoe & M. Shaffer, personal communication, April 24, 2018). Although this cost is steep, there may be economies of scale, such that community networks can employ coaches that serve multiple sites, so that a single program does not bear the full cost for a coach's training.

Using the MRS, I estimate that programs participating in MMCI on average spent $229.94 per teacher on PD-related expenses (e.g. training, coaching, substitute teachers, and planning time). The MRS denotes that the average program has 15 teachers, suggesting that aggregate costs for PD-related expenses are $3,449.12. These costs are intended to capture both materials associated with the MMCI coaching model, as well as paying for teachers' time to participate. Thus, the total costs associated with MMCI is estimated to be $7,449.12. Relative to an effectiveness estimate of 0.44, the

cost-effectiveness ratio is 0.059 increase in CLASS score for every $1,000 invested. However, this could be considered a lower-bound: if even just two programs split the costs of coach training, then the cost-effectiveness ratio increases to 0.081.

Equity. Heterogeneity analyses suggest that that MMCI may have had a stronger effect for programs serving communities with fewer Hispanic families ($p < 0.05$), which is categorized as an inequitable outcome. However, on all other variables examined (% Black non-Hispanic, median household income, and baseline CLASS scores), there were no signs of significant differences in effectiveness of MMCI.

Political feasibility. MMCI has been demonstrated to be received positively by program leaders and teachers alike (Early et al., 2017). This is likely because of several reasons. First, the MMCI model is less demanding of teachers' time, relative to PD programs such as MTP, where teachers must record themselves teaching, upload videos to their coach, and respond to prompts. Second, the group coaching model is anecdotally more aligned with how PD is typically delivered in ECE programs, and it enhances a program's professional culture through regular meetings and shared discussion for continuous improvement. Third, MMCI coaches are able to see variation in teaching practices among their group of teachers, which affords them greater flexibility in customizing feedback on a wider array of teaching examples in a way that one-on-one models may not allow. Finally, teachers report enjoying the content discussed in group coaching sessions (B. Hamre, personal communications, April 24, 2018).

Table 5. Assessment of MMCI policy option

	Effectiveness	Costs	Cost-Effectiveness	Equity	Feasibility
Making the Most of Classroom Interactions	0.44	$7,449.12	+0.059 in CLASS per $1,000 spent	3	5

Encourage teachers to participate in CLASS PreK reliability training

Cost-effectiveness. Teachstone offers CLASS PreK certification training for $900 per person, although this cost is projected to increase within the next several years (M. Shaffer & K. Kehoe, personal communication, April 27, 2018).

Using the MRS, I estimate that programs adopting CLASS PreK on average spent $236.78 per teacher on PD-related expenses. The MRS denotes that the average program has 15 teachers, suggesting that aggregate costs for PD-related expenses are $3,551.77. The MRS also denotes that adopting programs on average had two staff members participating in CLASS training.

The total costs associated with CLASS PreK reliability training totals $5,351.77. Relative to an effectiveness estimate of 0.24, the cost-effectiveness ratio is 0.045 increase in CLASS score for every $1,000 invested.

Equity. CLASS PreK reliability training was found to have greater effectiveness for programs serving communities with fewer Black non-Hispanic families ($p < 0.01$), which is categorized as an inequitable outcome. There was no evidence of differences in effectiveness of CLASS PreK reliability training on other variables.

Political feasibility. Unlike MMCI, which involves teachers but is delivered at a group setting, CLASS PreK reliability training is a per-teacher PD tool. Thus, CLASS reliability training targets effective teaching practices, but may not necessarily improve the professional environment of

teachers. Because CLASS observer trainings often occur at regional events over the span of several days, this option is most feasible if the LDE can request onsite trainings so that multiple teachers and staff members can attend at the same time, without burdening programs with the need for substitute teachers.

Table 6. Assessment of CLASS PreK reliability training policy option

	Effectiveness	Costs	Cost-Effectiveness	Equity	Feasibility
CLASS PreK observer reliability training	0.24	$5,351.77	+0.045 in CLASS per $1,000 spent	2	3

Ensure the majority of teachers have earned the Child Development Associate credential

Cost-effectiveness. The primary costs associated with hiring teachers with CDA is the increase in wages. Recall that the LDE is requiring all teachers to hold an *Early Childhood Ancillary Certificate* beginning July 2019. Teachers can earn the certificate either by holding a CDA, or by holding a technical diploma in an early childhood related field from a community college. Thus, I am interested in the increase in wage costs associated with paying teachers for a CDA instead of their previous degree. Descriptive statistics suggest that while the share of teachers with LA certification does not meaningfully differ across centers with and without 50% or more of its teachers with a CDA, the share of teachers with only a high-school degree differs by 39.3%.

Thus, I make the following assumptions in computing expected costs associated with this option. First, I assume that program leaders would respond to this policy option by shifting teachers away from the "high school diploma or equivalent" category toward the CDA credential. Second, I use responses from the MRS to estimate an average hourly wage rate of $9.72 for teachers holding a CDA, and $8.00 for teachers with a high school diploma or equivalent[6]. Third, respondents on average reported that 79% of their staff were employed full-time, and 21% were employed part-time. Fourth, I assume part-time as implying 20 hours per week, while full-time is 40 hours per week. Finally, respondents on average reported employing fifteen teachers (lead or assistant).

Based on these assumptions, the expected annual wages for CDA credential holders are $18,090.36, while the expected annual wages for teachers with high school diploma or equivalent are $14,891.83. Thus, for each teacher that obtains the CDA credential, the expected change in wages is $3,198.53. The average composition of CDA teachers was 23.2% among programs in the sample, suggesting that the average program would need to increase its composition of teachers with a CDA by 26.8%. Assuming 15 teachers in a program, this translates to roughly four teachers with increased wages, which totals $12,794.11. Finally, relative to an effectiveness of 0.16, my cost-effectiveness estimate is a 0.013 increase in CLASS score per $1000 invested.

Equity. There was no evidence of differences in effectiveness of CLASS PreK reliability training on any variables examined (% Black non-Hispanic, % Hispanic, median household income, baseline

[6] While I rely on the empirically derived estimate for this analysis, it is true that Louisiana's statewide minimum hourly wage rate was $7.25 per hour. If we assume measurement error and that teachers with high school diplomas or equivalent were in fact paid at minimum wage (as is commonly characterized in popular press), then the differential between these teachers and CDA-holders would increase, and the cost-effectiveness ratio would decrease. For this report, however, I rely on estimates from the data.

CLASS scores). Although these patterns do not necessarily promote equity as operationalized by my criteria, it is nevertheless encouraging that inequities are not further perpetuated.

Political feasibility. Feasibility of this requirement is rated highly in part because the LDE will require the Early Childhood Ancillary Certificate by July 2019, and although there are two pathways toward earning this credential, program leaders may be grappling with this decision anyway. While there may initially be concerns that ECE program leaders will want flexibility in their hiring practices and the extent to which they support their teachers in earning the credential, responses from the MRS suggest that programs on average hire about three teachers with a CDA for every one teacher with a technical diploma. Thus, I consider this option very feasible, in the sense that program leaders may already be making decisions that align with this policy option.

Table 7. Assessment of CDA policy option

	Effectiveness	Costs	Cost-Effectiveness	Equity	Feasibility
Child Development Associate credential	0.16	$12,794.11	+0.013 in CLASS per $1,000 spent	4	4

Outcomes Matrix

Below I summarize the policy analysis described above on each of the evaluative criteria. Sensitivity tests suggest that the final outcome of the analysis will not change unless (1) the weighting structure is changed to .2/.6/.2, or (2) the cost-effectiveness ratio for MMCI is rated as 3 *and* the weighting structure is adjusted to .3/.5/2; in both these scenarios, Options 2 and 4 would tie in average score. Assigning option 4 a feasibility score of 5 would increase its average score to 3.4

Table 8. Outcomes matrix for all policy options

	Cost-Effectiveness (40%) *Change in CLASS/$1,000*	Equity (40%)	Feasibility (20%)	Average Score
Option 1: Status Quo	---	2	---	---
Option 2: Implement Making the Most of Classroom Interactions	4 0.059 CLASS/$1000	3	5	3.8
Option 3: Encourage teachers to participate in CLASS PreK reliability training	3 0.045 CLASS/$1000	2	3	2.6
Option 4: Ensure the majority of teachers have earned the CDA credential	2 0.013 CLASS/$1000	4	4	3.2

RECOMMENDATION AND IMPLEMENTATION

In a policy landscape where programs are expected to respond to state-defined goals for quality, it is important for states to understand what responses they can employ to support programs' efforts toward improvement. **My recommendation is that the LDE support community networks' efforts in investing in coaching programs such as *Making the Most of Classroom Interactions*** (MMCI). Only 30% of programs reported participating in MMCI, yet my analysis reveals that this classroom support can be a particularly cost-effective lever toward quality improvement. MMCI further has the benefit of being scalable, in that community networks can hire coaches that can be employed across schools; and of being well received by multiple stakeholders, particularly among teachers.

Two primary challenges face the LDE in implementing MMCI at scale. First, MMCI has high start-up costs, with $4,000 for training per coach. To distribute costs and reduce budgetary burden, the LDE might recommend that parishes either roll out implementation of MMCI or hire coaches that serve multiple sites in the pilot year. Second, my results tentatively suggest that MMCI may have been less effective for programs serving communities with a greater proportion of Hispanic families. It is difficult to interpret this finding, given that I do not have measures of implementation; nevertheless, as the state implements this policy option at scale, it should ensure that the group coaching content considers the wide range of classroom environments, and that it adequately supports teachers to lead diverse classrooms.

The findings from this report have substantive implications for both policymakers and researchers. For the LDE, it provides quantifiable evidence on which levers effectively support quality classroom interactions important for children's development. This information can help program leaders target their improvement efforts toward efficient quality investments, and it can help the LDE target resources and supports to ensure programs have equitable access to improvement levers. Yet, more work needs to be done to better understand the ways in which states can both operationalize quality and support programs in their improvement efforts. As the LDE begins to implement policy investments at scale, they should consider a future analysis that not only examines quality outcomes on the longer term, to understand the extent to which improvements can be sustained at scale; but also includes more precise measures and measures of implementation to better identify the effects of specific levers on quality outcomes.

Virtually every state implements a QRIS, and although the design of Louisiana's QRIS is certainly unique, the findings here have important implications for policymakers at large as well. Indeed, the research literature to date has not carefully examined program responses to QRIS, and this report is among the first studies to understand improvement in classroom quality at scale and outside a research-controlled environment. As policymakers identify quality components integral to their state' QRIS design and explore ways to incentivize programs to improve on these components, they should consider supports that are closely aligned with these measures.

References

Ansari, A., & Winsler, A. (2014). Montessori public school pre-k programs and the school readiness of low-income black and latino children. *Journal of Educational Psychology*, *106*(4), 1066–1079. http://doi.org/10.1037/a0036799

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, *131*(3), 1415–1453. http://doi.org/10.1093/qje/qjw016.Advance

Bassok, D., Fitzpatrick, M., Greenberg, E., & Loeb, S. (2016). Within- and between-sector quality differences in early childhood education and care. *Child Development*, *87*(5), 1627–1645. http://doi.org/10.1111/cdev.12551

Bassok, D., & Galdo, E. (2016). Inequality in preschool quality? Community-level disparities in access to high-quality learning environments. *Early Education and Development*, *27*(1), 128–144. http://doi.org/10.1080/10409289.2015.1057463

Burchinal, M. R., Roberts, J. E., Rhodus Riggins, J., Zeisel, S. A., Neebe, E., & Bryant, D. (2000). Relating Quality of Center-Based Child Care to Early Cognitive and Language Development Longitudinally Published by : Wiley on behalf of the Society for Research in Child Development Stable URL : http://www.jstor.org/stable/1131994 REFERENCES Linked refer. *Child Development*, *71*(2), 339–357.

Burchinal, M. R., Vandergrift, N., Pianta, R. C., & Mashburn, A. J. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*(2), 166–176. http://doi.org/10.1016/j.ecresq.2009.10.004

Campbell, F. A., Pungello, E. P., Burchinal, M. R., Kainz, K., Pan, Y., Wasik, B. H., … Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An abecedarian project follow-up. *Developmental Psychology*, *48*(4), 1033–1043. http://doi.org/10.1037/a0026644

Cannon, J. S., Zellman, G. L., Karoly, L. A., & Schwartz, H. L. (2017). Quality Rating and Improvement Systems for early care and education programs: Making the second generation better. Retrieved from https://www.rand.org/pubs/perspectives/PE235.html

Clements, D. H. (2007). Curriculum Research: Toward a Framework for "Research-Based Curricula." *Journal for Research in Mathematics Education*, *38*(1), 35–70.

Connors, M. C. (2016). Creating cultures of learning: A theoretical model of effective early care and education policy. *Early Childhood Research Quarterly*, *36*, 32–45. http://doi.org/10.1016/j.ecresq.2015.12.005

Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. A., & Jones, D. (2009). Fostering High-Quality Teaching With an Enriched Curriculum and Professional Development Support:The Head Start REDI Program. *American Educational Research Journal*, *46*(2), 567–597. http://doi.org/10.3102/0002831208328089

Duncan, G. J., Jenkins, J. M., Auger, A., Burchinal, M. R., Domina, T., & Bitler, M. (2015). *Boosting school readiness with preschool curricula*. *Irvine Network on Interventions in Development*. http://doi.org/10.1017/CBO9781107415324.004

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., … Vandergrift, N. (2007). Teachers' Education, Classroom Quality, and Young Children' s Academic Skills: Results from Seven Studies of Preschool Programs. *Child Development*, *78*(2), 558–580.

Early, D. M., Maxwell, K. L., Ponder, B. B., & Pan, Y. (2017). Improving teacher-child interactions: A randomized control trial of Making the Most of Classroom Interactions and My Teaching Partner professional development models. *Early Childhood Research Quarterly*, *38*, 57–70. http://doi.org/10.1016/j.ecresq.2016.08.005

Fukkink, R. G., & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly*, *22*(3), 294–311. http://doi.org/10.1016/j.ecresq.2007.04.005

Hamre, B. K. (2014). Teachers' Daily Interactions With Children: An Essential Ingredient in Effective Early Childhood Programs. *Child Development Perspectives*, *8*(4), 223–230. http://doi.org/10.1111/cdep.12090

Hamre, B. K., Hatfield, B. E., Pianta, R. C., & Jamil, F. (2014). Evidence for General and Domain-Specific Elements of Teacher-Child Interactions: Associations With Preschool Children's Development. *Child Development*, *85*(3), 1257–1274. http://doi.org/10.1111/cdev.12184

Hamre, B. K., Partee, A., & Mulcahy, C. (2017). Enhancing the Impact of Professional Development in the Context of Preschool Expansion. *AERA Open*, *3*(4), 233285841773368. http://doi.org/10.1177/2332858417733686

Hamre, B. K., Pianta, R. C., Burchinal, M. R., Field, S., LoCasale-Crouch, J., Downer, J. T., … Scott-Little, C. (2012). A Course on Effective Teacher-Child Interactions: Effects on Teacher Beliefs, Knowledge, and Observed Practice. *American Educational Research Journal*, *49*(1), 88–123. http://doi.org/10.3102/0002831211434596

Hatfield, B. E., Hestenes, L. L., Kintner-Duffy, V. L., & O'Brien, M. (2013). Classroom Emotional Support predicts differences in preschool children's cortisol and alpha-amylase levels. *Early Childhood Research Quarterly*, *28*(2), 347–356. http://doi.org/10.1016/j.ecresq.2012.08.001

Hatfield, B. E., Lower, J. K., Cassidy, D. J., & Faldowski, R. A. (2015). Inequities in access to quality early care and education: Associations with funding and community context. *Early Childhood Research Quarterly*, *30*(PB), 316–326. http://doi.org/10.1016/j.ecresq.2014.01.001

Hotz, V. J., & Xiao, M. (2011). The Impact of Regulations on the Supply and Quality of Care in Child Care Markets. *American Economic Review*, *101*(5), 1775–1805. http://doi.org/10.1257/aer.101.5.1775.The

Knapp, M. S. (2003). Professional Development as a Policy Pathway. *Review of Research in Education*, *27*(1), 109–157.

Kraft, M. A., Blazar, D., & Hogan, D. (2017). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence.

Lin, Y. C., & Magnuson, K. A. (2018). Classroom quality and children's academic skills in child care centers: Understanding the role of teacher qualifications. *Early Childhood Research Quarterly*, *42*(April 2016), 215–227. http://doi.org/10.1016/j.ecresq.2017.10.003

Mashburn, A. J., Pianta, R. C., Barbarin, O. A., Bryant, D., Hamre, B. K., Downer, J. T., … Early, D. M. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–749.

Maxwell, K. L., Field, C. C., & Clifford, R. M. (2006). Defining and measuring professional development in early childhood research. In M. J. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in Early Childhood Development* (pp. 21–48). Baltimore, MD: Paul H Brookes Publishing.

Payne, A. L. (2011). *Strong licensing: the foundation for a quality early care and education system*. Lexington.

Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., … Weiland, C. (2017). Puzzling It Out: The Current State of Scientific Knowledge on Pre-Kindergarten Effects A Consensus Statement. Retrieved from https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement_final.pdf

Phillips, D. A., McCartney, K., & Sussman, A. (2008). Child Care and Early Development. In K. McCartney & D. A. Phillips (Eds.), *Blackwell Handbook of Early Childhood Development* (pp. 471–489). Oxford: Blackwell Publishing.

Pianta, R. C., Barnett, W. S., Burchinal, M., & Thornburg, K. R. (2009). The effects of preschool education: What we know, how public policy is or is not aligned with the evidence base, and what we need to know. *Psychological Science in the Public Interest, Supplement*, *10*(2), 49–88.

http://doi.org/10.1177/1529100610381908
Pianta, R. C., Howes, C., Burchinal, M. R., Bryant, D., Clifford, R., Early, D. M., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159. http://doi.org/10.1207/s1532480xads0903
Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). CLASS: Classroom Assessment Scoring System. *Assessment*. Paul H Brookes Publishing.
Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Education and Development*, *23*, 431–451. http://doi.org/10.1080/10409289.2015.1047711
Sabol, T. J., & Pianta, R. C. (2014). Do standard measures of peschool quality used in statewide policy predict srchool readiness? *Education Finance and Policy*, *9*(2), 116–164. http://doi.org/10.1162/EDFP_a_00127
Schachter, R. E. (2015). An analytic study of the professional development research in early childhood education. *Early Education and Development*, *26*(8), 1057–1085. http://doi.org/10.1080/10409289.2015.1009335
Schweinhart, L., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry preschool study through age 40*. Ypsilanti: High Scope Press.
Sheridan, S. M., Edwards, C. P., Marvin, C. A., & Knoche, L. L. (2009). Professional development in early childhood programs: Process issues and research needs. *Early Education and Development*, *20*(3), 377–401. http://doi.org/10.1080/10409280802582795
Zellman, G. L., & Perlman, M. (2010). *Child-care Quality Rating and Improvement Systems in five pioneer states: Implementation issues and lessons learned*. http://doi.org/RAND

Appendix A. Performance Profiles

2016-2017 EARLY CHILDHOOD SITE **PERFORMANCE PROFILE GUIDE**

Research shows a strong connection between the quality of early childhood experiences and future academic success. The **Performance Profile** helps parents understand what children are likely to experience at an Early Childhood Program. These are only generated for publicly-funded sites, which includes Head Starts, public PreK, nonpublic PreK (NSECD), or child care centers receiving CCAP.

MEASURES OF CLASSROOM QUALITY

Louisiana's classrooms are observed using *CLASS*™, a nationally recognized tool. These in-depth measures show how well classrooms support children's growth and development.

PREK (3-4 years)		**EMOTIONAL SUPPORT** measures the responsiveness and sensitivity of the classroom to children's emotions, which helps develop their ability to interact with others.
		CLASSROOM ORGANIZATION measures how a classroom is organized to manage children's attention, time, and behavior.
		INSTRUCTIONAL SUPPORT measures how classroom interactions and experiences help children develop language and learn new concepts.
TODDLER (1-2 years)		**EMOTIONAL & BEHAVIORAL SUPPORT** measures the responsiveness and sensitivity of the classroom to children's emotions, which helps develop their ability to interact with others and how a classroom is organized to manage children's attention, time, and behavior.
		ENGAGED SUPPORT FOR LEARNING measures how classroom interactions and experiences help children develop language and learn new concepts.

CLASSROOM SYMBOLS

This represents the areas that address children's emotional needs.

This represents the areas that address routines and organization to help children learn.

This represents the areas that develop children's language and critical thinking skills.

USE OF BEST PRACTICES

Stars help parents understand the practices in their children's classrooms. This self-reported information is not included in the rating.

Assessment This tells you how many publicly-funded children are assessed.	★★★ Most			★★☆ Some			★☆☆ Few		None
Children per Teacher This measures the ratio of children per teacher. The smaller the ratio, the more individual attention children can receive. (Children:Teacher:Maximum Group Size)	Age	Ratio	Group	Age	Ratio	Group	Age	Ratio	Site did not report ratio information.
	0-1	4:1	8	0-1	4:1	8	0-1	5:1	
	1-2	4:1	8	1-2	6:1	12	1-2	7:1	
	2-3	6:1	12	2-3	8:1	16	2-3	11:1	
	3-4	8:1	16	3-4	10:1	20	3-4	13:1	
	4-5	10:1	20	4-5	12:1	24	4-5	15:1	
Curriculum Quality This measures how the site's curriculum compares with the state's indicators of quality. Using high-quality activities promotes kindergarten readiness.	Meets ALL indicators of a high-quality curriculum.			Meets SOME indicators of a high-quality curriculum.			Meets FEW indicators of a high-quality curriculum.		Site does not report using any curriculum or is using one that has not been rated.

TEACHER DEGREES & CERTIFICATION

Louisiana teachers need excellent training and ongoing professional development to provide children experiences they need to develop. Degree requirements vary by program type and site director's preferences*.

No Degree	The instructor holds no degree. (0 credit hours/some credit hours but no degree or certificate)
Child Development Associates	The instructor holds a professional certificate catered to child development. (120 clock hours)
Associates	The instructor holds an Associate's degree from a college or university. (60+ credit hours)
Bachelors	The instructor holds at least a Bachelor's degree from a college or university. (120+ credit hours)
Masters +	The instructor holds a Master's degree or higher from a college or university. (180+ credit hours)

* The number of teachers with certifications indicates the percent of teachers with an applicable teaching certificate. Teaching certificate requirements vary by type.

Appendix B. Market-Rate Survey Questions

Note: Only items used in the analysis are reproduced below. Question numbers are preserved from the hardcopy version of the survey.

Teacher Hiring: Wages, Employment, and Education Level

Please write in each box the number of full-time and part-time teachers for each question below.	# of full-time Lead Teachers	# of part-time Lead Teachers	# of full-time Assistant Teachers	# of part-time Assistant Teachers
32. How many current teachers have been employed in this center/program for 3 or more years?				
33. How many current teachers have at least 3 years of experience in any licensed child care center, Early Head Start or Head Start program, or pre-K program?				
34. How many teachers left your employment from January 1, 2016 to December 31, 2016?				
35. How many new teachers did you hire from January 1, 2016 to December 31, 2016?				

36. Please write in each box the number of directors/teachers whose highest education level is the degree or education level listed, and also list starting hourly wage and average hourly wage for teachers in each education level.

Count each person only once for his/her highest degree. For example, if a teacher has both a bachelor's degree and a master's degree, you would count that teacher in the master's degree row but not in the bachelor's degree row.	# of Directors or Assistant Directors	# of Lead Teachers	# of Assistant Teachers	Starting Hourly Wage	Average Hourly Wage
Ph.D.					
Master's degree					
Bachelor's degree					
Associate degree					
Child Development Associate (CDA)					
Vocational/technical school diploma/certificate					
High school diploma or Equivalent					
Child Care Career Diploma					
None of the above					

Professional Development

15. How much did you spend for on professional development for teachers and directors in each of the following categories from January 1, 2016 to December 31, 2016?

	Amount Spent		**Amount Spent**
Training		Substitutes for teachers being trained	
Conferences		Stipends (do not include bonuses)	
Coaching		Paid time for planning	

16. Have you (staff including teachers, owners and directors) participated in the following? If Yes, how many staff attended?

Training on curriculum (e.g. Creative Curriculum, or Frog Street)	No ❑	Yes ❑	Number of staff: _____
Making the Most of Classroom Interactions (MMCI)	No ❑	Yes ❑	Number of staff: _____
My Teaching Partner (MTP)	No ❑	Yes ❑	Number of staff: _____
MyTeachStone	No ❑	Yes ❑	Number of staff: _____
CLASS*TM* PreK observer reliability training	No ❑	Yes ❑	Number of staff: _____
CLASS*TM* Toddler observer reliability training	No ❑	Yes ❑	Number of staff: _____
GOLD training	No ❑	Yes ❑	Number of staff: _____

Curricular Materials

13. How much did you spend to purchase curriculum materials from January 1, 2016 to December 31, 2016?

_____

14. If purchased, what curricular materials did you purchase?

_____

Appendix C. Sample Selection and Missing Data Analysis

This appendix serves two purposes: Understanding missingness in MRS responses, and describing the analytic technique used in this report to deal with missing data.

Matching Procedure and Match Rates

Louisiana Child Care Market Rate Survey. Responses were linked to administrative data using a combination of programs' operating license number, program name, and site code (a state-assigned id variable). However, only about 42 percent of programs provided financial data in their responses. These programs were contacted again and invited to respond to these items. Thus, as of March 29th, 2018, 67 percent of all child care programs (Type I, II, and III) had responded to the survey, or about 1,261 responses.

Of these responses, 46.9% matched to administrative data, 34.1% could not be linked, and 19.0% were duplicate responses (e.g., supplementing responses with financial data, accidental openings of the online survey). Some responses could not be linked to administrative data due to lack of identifiable information. However, the primary reason for unmatched survey responses was because although all center-based providers were invited to participate, only publicly-funded programs received CLASS scores. In addition, school-based programs in Louisiana do not receive a license number. Thus, my data primarily represent Type III programs, which are my population of interest given that non-school-based programs may face the most challenges toward quality improvement.

LDE administrative data. In the 2016-17 school year, 1,502 publicly-funded programs received ratings from the LDE; ratings from schools in Livingston Parish were not provided due to flooding. Of these, 861 had a license number available for merging. These programs primarily comprised Early/Head Start programs (N=190) and child care centers (N=660), given that school-based programs (N=641) do not receive license numbers from the state; an additional 11 programs did not have a listed program type.

Of the remaining programs, 509 programs (59.1%) could be matched to MRS data. Twenty-seven of these programs (5.3%) did not have CLASS data from the 2015-16 school year, meaning that *changes* in quality outcomes over time could only be computed for 482 programs. Missingness and descriptive statistics are provided on the next page.

Table C1. Missingness and descriptive statistics, all variables

	Overall				Child Care				Early/Head Start			
	% valid	valid n	Mean	SD	% valid	valid n	Mean	SD	% valid	valid n	Mean	SD
Panel A: Match rates to MRS												
No. of sites with two years of CLASS data EXCLUDING school-based		797				610				187		
Match rate to MRS EXCLUDING school-based	60.5%	482			66.6%	406			40.6%	76		
Panel B: Response rates to survey items												
Tier 1 curricula	39.4%	190	0.71	0.45	44.1%	179	0.70	0.46	14.5%	11	0.91	0.30
Curricula-related training	78.2%	377	0.77	0.42	79.1%	321	0.76	0.43	73.7%	56	0.86	0.35
MMCI	73.9%	356	0.30	0.46	75.6%	307	0.31	0.46	64.5%	49	0.27	0.45
MTP	73.7%	355	0.15	0.36	74.9%	304	0.14	0.35	67.1%	51	0.20	0.40
MyTeachStone	76.1%	367	0.74	0.44	77.3%	314	0.73	0.44	69.7%	53	0.75	0.43
CLASS PreK obs training	77.2%	372	0.75	0.43	78.1%	317	0.75	0.43	72.4%	55	0.75	0.44
CLASS Toddler obs training	75.7%	365	0.69	0.46	77.1%	313	0.72	0.45	68.4%	52	0.50	0.50
GOLD training	77.6%	374	0.95	0.22	78.8%	320	0.95	0.22	71.1%	54	0.96	0.19
50% or more with CDA	94.6%	456	0.22	0.42	93.6%	380	0.25	0.43	100.0%	76	0.08	0.27

Notes. N represents total number of programs in administrative data. Match % describes how many of the N programs matched to market rate survey data. The matched sample contains n = 76 Early/Head Start programs and n = 406 child care centers.

What Explains Missingness?

I explore two types of missingness that may introduce bias: (1) likelihood of matching with MRS responses, and (2) likelihood of missing data on specific survey items, conditional on having responded. For both types of missingness, I run logistic regressions to determine whether (a) baseline CLASS scores, (b) program type, or (c) teacher composition predict missingness.

What types of programs were more likely to take the market rate survey? Table C2 summarizes logistic regressions for administrative variables predicting missingness on MRS. A clear pattern that emerges is that Head Start programs were more likely than child care centers to have missing data, although it is not immediately obvious why these response rates vary. Thus, the subsequent rows of Table C2 control for this variable, to understand other determinants of missing data after adjusting for program type.

After adjusting for program type, lower quality programs seemed to have higher odds of having missing data. The coefficient is statistically significant for the overall score. Although initially teacher composition explained missingness, these effects washed out after adjusting for program type; programs with greater shares of teachers with a CDA were slightly more likely to respond to

Table C2. Summary of logistic regression analysis for administrative variables predicting missingness on market rate survey

Predictor	B	Robust SE	Odds-Ratio
Program Type			
1 = Head Start	1.060 ***	0.166	2.887
2015-16 CLASS*			
Overall	-0.303 **	0.125	0.739
Percent of Teachers with…*			
Master's	0.010	0.007	1.010
Bachelor's	0.003	0.003	1.003
Associate	-0.002	0.003	0.998
CDA	-0.005 †	0.003	0.995
LA Cert.	-0.001	0.002	0.999
HS equiv.	0.002	0.002	1.002

Notes. Outcome variable in all logistic regressions is whether programs in administrative data did not match (i.e., 1 = missingness) to market rate survey responses, The Odds-Ratio is computed through exponentiation of B (i.e., e^B). Odds-Ratio greater than 1 indicates a positive relationship between the predictor variable and missingness; a ratio less than 1 denotes a negative association. Significance: † p < 0.1; * p < 0.05; ** p < 0.01; *** p < 0.001

*Logistic regressions control for program type.

the survey. Overall, one out of 10 regressions were significant at the 5% level, which is more frequently than what we might expect simply by chance.

What types of respondents were more likely to omit classroom support questions? Tables C3-C4 summarize logistic regressions predicting missingness for specific items on the MRS, used in this analysis. Across all classroom support measures, program type strongly predicted missing data, with Head Start having the greatest odds for missingness. Programs with smaller proportions of teachers with Associate degrees, as well as greater shares of teachers without certification or a college degree, also had greater odds of having missing data.

Addressing Missing Data

The results above are suggestive that missingness may not be completely at random: program type, baseline quality, and teacher composition may be predictive of missing data. One worry with simply dropping programs with missing data is that missingness may be a confounding variable that influences both market rate responses and quality outcomes, which therefore biases any estimations I conduct. In other words, we may not be describing a sample that is representative of the larger population of ECE programs in Louisiana.

The approach I employ is a dummy variable approach, in which for each MRS variable I include a dichotomous variable that is equal to 1 if the survey item is missing, and zero otherwise. Missingness in the original variable is then coded as zero. Thus, instead of one dichotomous variable representing a two-factor variable with missing data, I instead have a three-factor variable represented by two dummy variables ("Yes" and "missing," with "No" as the reference category).

Table C3. Logistic regressions of program and parish characteristics predicting missingness on curricula and teacher preparation variables, conditional on matched sample

	Tier 1 Curricula	Curricula-related training	CDA	LA Cert.
2015-16 CLASS	-0.14	0.10	-1.10***	-0.76**
	(0.16)	(0.18)	(0.32)	(0.27)
1 = Head Start	1.40*	0.44		
	(0.67)	(0.59)		
% BA+	0.00	-0.01†		
	(0.01)	(0.01)		
1 = two classrooms	-1.10**	-0.70*		
	(0.38)	(0.28)		
1 = three classrooms	-0.97*	-0.80*		
	(0.47)	(0.39)		
% unemployed above state median	0.06	0.14	-0.20	-1.08
	(0.31)	(0.28)	(0.73)	(0.66)
% maternal labor force above state median	-0.05	-0.32	0.91	0.91†
	(0.34)	(0.36)	(0.82)	(0.53)
% Black, non-Hispanic above state median	-0.19	-0.10	0.92†	0.31
	(0.19)	(0.28)	(0.55)	(0.54)
% Hispanic above state median	-0.53*	0.49	1.14*	1.26**
	(0.23)	(0.38)	(0.57)	(0.48)
Median household income above state median	-0.02	-0.39	0.31	-1.48*
	(0.35)	(0.33)	(0.70)	(0.67)
No. of children under 5 above state median	-0.22	-0.33	-1.19	-0.71
	(0.37)	(0.47)	(0.72)	(0.67)
Constant	2.64**	-0.55	0.47	1.25
	(1.00)	(1.03)	(2.01)	(1.70)
Observations	481	481	481	481

What this approach allows is to preserve the full sample size, while also allowing me to assess the association between the missing data indicator and outcomes of interest. In other words, constructing the variable and including it in the regression estimations allow me to 'adjust' my estimates on missingness. One limitation to this approach, however, is that it is difficult to implement with continuous predictors.

There are other methods for addressing missing data, such as multiple imputation and maximum likelihood estimation, though these are more complex methods. In future analyses, I will examine the sensitivity of my main findings to different missing data techniques.

Table C4. Logistic regressions of program and parish characteristics predicting missingness on PD variables, conditional on matched sample

	MMCI	MTP	MyTeachStone	CLASS PreK reliability training	CLASS Toddler reliability training	GOLD training
2015-16 CLASS	-0.01	0.04	0.02	-0.05	-0.07	0.11
	(0.19)	(0.20)	(0.21)	(0.21)	(0.20)	(0.18)
1 = Head Start	0.68	0.52	0.68	0.65	0.74	0.69
	(0.46)	(0.51)	(0.53)	(0.55)	(0.50)	(0.52)
% BA+	-0.01	-0.01	-0.01*	-0.01*	-0.01†	-0.01*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
1 = two classrooms	-1.00***	-0.65*	-0.85***	-0.78**	-0.54*	-0.61*
	(0.22)	(0.25)	(0.23)	(0.25)	(0.27)	(0.27)
1 = three classrooms	-1.18**	-1.09**	-1.06**	-0.73†	-1.01*	-0.97*
	(0.38)	(0.38)	(0.36)	(0.42)	(0.42)	(0.46)
% unemployed above state median	0.04	0.04	-0.05	0.03	0.06	0.04
	(0.29)	(0.26)	(0.28)	(0.27)	(0.26)	(0.26)
% maternal labor force above state median	-0.15	-0.10	-0.25	-0.16	-0.08	-0.22
	(0.31)	(0.33)	(0.30)	(0.36)	(0.34)	(0.34)
% Black, non-Hispanic above state median	-0.10	-0.14	-0.13	-0.15	-0.10	-0.22
	(0.23)	(0.27)	(0.25)	(0.26)	(0.31)	(0.28)
% Hispanic above state median	0.42	0.18	0.49	0.53	0.58	0.65†
	(0.34)	(0.37)	(0.33)	(0.36)	(0.40)	(0.36)
Median household income above state median	-0.36	-0.19	-0.34	-0.39	-0.39	-0.50
	(0.30)	(0.32)	(0.30)	(0.32)	(0.35)	(0.31)
No. of children under 5 above state median	-0.26	-0.26	-0.47	-0.49	-0.35	-0.46
	(0.48)	(0.48)	(0.42)	(0.47)	(0.48)	(0.46)
Constant	0.27	-0.15	0.22	0.32	0.00	-0.54
	(1.00)	(1.08)	(1.11)	(1.19)	(1.14)	(1.00)
Observations	481	481	481	481	481	481

Appendix D. Access and Effectiveness Results

Explanation of Empirical Strategy

The main report describes the intuition for fitting a model with prior CLASS scores, as well as program and parish characteristics. While this model addresses the concerns raised, in practice it is often insightful to introduce sets of covariates incrementally, to assess the sensitivity of the coefficient of interest to various variables that may be potentially confounding.

Consider again two programs receiving comparable CLASS scores in the 2015-16 school year. Based on this score, each program leader is faced with decisions on how to marshal and reallocate resources toward quality improvement, which is then reflected in their 2016-17 CLASS score. Thus, part of the 2016-17 score is determined by programs' previous score, and the estimate we are interested in is: above and beyond these determinants, is the decision to invest in a given classroom support associated with higher quality outcomes?

To answer this question, I can start by specifying a basic regression specification (Model 1):

$$CLASS_{ipt} = \alpha + \beta_1(Policy = 1)_{ipt} + \beta_2 CLASS_{ip(t-1)} + \bar{\boldsymbol{M}}_i + \epsilon_{ipt} \quad \text{(Model 1)}$$

where $(Policy = 1)_{ipt}$ is a program-level measure for whether a program *i* adopted the policy option, $CLASS_{ip(t-1)}$ and $CLASS_{ipt}$ represent the 2015-16 and 2016-17 CLASS scores for program *i* in parish *p*, respectively, and $\bar{\boldsymbol{M}}_i$ is a series of variables that indicate missingness in the original survey data (see Appendix C for a detailed explanation).

The parameter β_1 is our coefficient of interest, and represents the expected difference in average CLASS score in 2016-17 between those that invested in the option $(Policy = 1)_{ipt}$ and those that did not $(Policy = 0)_{ipt}$, adjusting for what CLASS scores programs received in the year prior.[7]

Our primary concern with Model 1 was that program type and program characteristics are related to program quality and may also be related to decisions to use or not use classroom supports. Ignoring parish-level influences for now, I can estimate a second model that includes these program-level variables (Model 2):

$$CLASS_{ipt} = \alpha + \beta_1(Policy = 1)_{ipt} + \beta_2 CLASS_{ip(t-1)} + \bar{\boldsymbol{M}}_i + \sum_{x \in \bar{\boldsymbol{X}}_i} \beta_3 x + \epsilon_{ipt} \quad \text{(Model 2)}$$

where $\bar{\boldsymbol{X}}_i$ is a series of program-level characteristics (e.g. Head Start, % of teachers with at least a Bachelor's degree), and all other variables retain their meaning. The parameter β_1 continues to be our coefficient of interest and is now adjusted for program characteristics.

While Model 2 is an improvement to Model 1 in that it adjusts for potentially confounding, program-level factors, it still does not account for the possibility that community characteristics—above and beyond programs themselves—may influence the kinds of decisions programs tend to make. Thus, I specified the primary model used in the main report:

$$CLASS_{ipt} = \alpha + \beta_1(Policy = 1)_{ipt} + \beta_2 CLASS_{ip(t-1)} + \bar{\boldsymbol{M}}_i + \sum_{x \in \bar{\boldsymbol{X}}_i} \beta_3 x + \sum_{p \in \bar{\boldsymbol{P}}_p} \beta_4 p + \epsilon_{ipt} \quad \text{(Model 3)}$$

where $\bar{\boldsymbol{P}}_p$ is a series of parish-level characteristics, described in Table 4, and all other variables retain their meaning. The parameter β_1 is now interpreted as the expected difference in CLASS

[7] Recall that standard errors are clustered at the parish level.

score between adopting and non-adopting programs, after adjusting for program and parish characteristics.

A final note: a fourth model I could run is a parish fixed effects model, which examines whether policy options are effective in improving CLASS outcomes even when comparing only programs located within the same parish. This approach has stronger internal validity than Model 3: While Model 3 requires parish-level confounding factors to be observed and thus adjusted for, the fixed effects approach adjusts for both *observed* and *unobserved* parish-level influences, so long as they are constant within parishes.

However, this approach requires within-parish variation in programs' decisions to use classroom supports. A fixed effects specification is difficult in the present context because (1) the majority of parishes had fewer than 10 programs with valid data, ranging from 64% to 74% across variables; and (2) many parishes did not have within-parish variation to identify estimates (in other words, within 42%-53% of parishes depending on the classroom support variable, all programs either adopted or did not adopt a classroom support variable). For these reasons, I prefer Model 3.

Main Results

This appendix presents effectiveness results from all Models described above, for all classroom support variables considered in this report. The main pattern I look for across these tables is whether the magnitude of the estimate changes substantively, and if the estimate retains its statistical and practical significance. If the estimates are sensitive from Model 1 to Model 2, then I would conclude that program characteristics are important in explaining quality outcomes. Similarly, if estimates change meaningfully in Model 3, then I would conclude community characteristics are important in explaining quality outcomes. Additionally, this appendix presents access results for all classroom variables.

Table D1. Logistic regressions of program and parish characteristics predicting use of curricula and teacher preparation supports, conditional on valid MRS responses

	Tier 1 Curricula	CDA
2015-16 CLASS	-0.34	0.95***
	(0.26)	(0.22)
1 = Head Start	1.30	-0.80
	(0.90)	(0.61)
% BA+	0.01	-0.03***
	(0.01)	(0.01)
% unemployed above state median	0.69	0.44
	(0.54)	(0.44)
% maternal labor force above state median	0.11	1.20**
	(0.52)	(0.43)
% Black, non-Hispanic above state median	-1.06**	0.10
	(0.36)	(0.34)
% Hispanic above state median	-0.49	-0.21
	(0.55)	(0.49)
Median household income above state median	0.02	0.29
	(0.62)	(0.56)
No. of children under 5 above state median	1.10	-0.40
	(0.99)	(0.56)
Constant	1.89	-6.10***
	(1.76)	(1.17)
Observations	189	481

Table D2. Logistic regressions of program and parish characteristics predicting use of PD supports, conditional on valid MRS responses

	MMCI	MTP	MyTeachStone	CLASS PreK reliability training	CLASS Toddler reliability training	GOLD training
2015-16 CLASS	-0.02	-0.02	0.52*	0.45*	0.49**	0.10
	(0.23)	(0.28)	(0.25)	(0.19)	(0.17)	(0.39)
1 = Head Start	-0.28	0.40	0.07	0.24	-0.63	1.38†
	(0.54)	(0.81)	(0.58)	(0.53)	(0.44)	(0.84)
% BA+	0.00	0.01	0.01	-0.01*	-0.01**	-0.03***
	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)
% unemployed above state median	-0.33	-1.17†	0.43	0.11	-0.59*	0.70
	(0.45)	(0.62)	(0.35)	(0.40)	(0.29)	(0.95)
% maternal labor force above state median	-0.50	1.14*	0.08	-0.06	-0.15	-0.26
	(0.48)	(0.55)	(0.32)	(0.45)	(0.23)	(0.72)
% Black, non-Hispanic above state median	0.45	-0.05	0.43	0.03	-0.11	0.64
	(0.29)	(0.31)	(0.29)	(0.34)	(0.21)	(0.57)
% Hispanic above state median	0.12	0.94†	1.03**	-0.15	0.54	-0.12
	(0.39)	(0.53)	(0.38)	(0.30)	(0.38)	(0.58)
Median household income above state median	-0.01	-0.95	-0.15	0.69†	-0.75†	0.71
	(0.45)	(0.71)	(0.37)	(0.36)	(0.42)	(0.90)
No. of children under 5 above state median	-0.52	0.13	-0.30	-0.02	-0.57	-0.77
	(0.53)	(0.92)	(0.50)	(0.45)	(0.43)	(0.86)
Constant	-0.09	-2.70	-2.15†	-1.02	-0.03	3.07
	(1.21)	(1.67)	(1.22)	(1.00)	(0.91)	(2.19)
Observations	355	354	366	371	364	373

Table D3. Regression analyses of curricula variables predicting 2016-17 CLASS scores, adjusting for program and parish covariates

	Model 1	Model 2	Model 3
Tier 1 Curricula			
Yes	0.19	0.17	0.19
	(0.17)	(0.17)	(0.19)
Curricula-related Training			
Yes	0.11	0.10	0.14
	(0.12)	(0.12)	(0.13)
Tier 1 x Training			
Yes#Yes	-0.11	-0.11	-0.15
	(0.17)	(0.18)	(0.19)
Program Characteristics			
2015-16 CLASS	0.57***	0.56***	0.56***
	(0.05)	(0.05)	(0.05)
1 = Head Start		-0.01	0.00
		(0.10)	(0.09)
% BA+		0.00	0.00
		(0.00)	(0.00)
1 = two classrooms		0.18†	0.27*
		(0.10)	(0.13)
1 = three classrooms		0.23†	0.26†
		(0.12)	(0.15)
Parish Characteristics			
% unemployed above state median			0.08
			(0.09)
% maternal labor force above state median			-0.04
			(0.08)
% Black, non-Hispanic above state median			-0.04
			(0.08)
% Hispanic above state median			0.02
			(0.08)
Median household income above state median			0.28***
			(0.08)
No. of children under 5 above state median			0.10
			(0.10)
Constant	1.98***	1.83***	1.44***
	(0.21)	(0.21)	(0.25)
Observations	481	481	481
Adjusted R-squared	0.345	0.345	0.384

Table D4. Regression analyses of teacher preparation variables predicting 2016-17 CLASS scores, adjusting for program and parish covariates

	Model 1	Model 2	Model 3
>=50% with CDA			
Yes	0.16†	0.16†	0.16*
	(0.09)	(0.09)	(0.08)
>=50% with LA Cert.			
Yes	0.05	0.04	0.06
	(0.07)	(0.07)	(0.06)
Tier 1 x Training			
Yes#Yes	-0.05	-0.03	-0.05
	(0.13)	(0.13)	(0.13)
Program Characteristics			
2015-16 CLASS	0.53***	0.53***	0.53***
	(0.06)	(0.06)	(0.06)
1 = Head Start		-0.02	-0.00
		(0.11)	(0.09)
% BA+		0.00	0.00
		(0.00)	(0.00)
1 = two classrooms		0.19†	0.27*
		(0.11)	(0.13)
1 = three classrooms		0.24†	0.27†
		(0.12)	(0.15)
Parish Characteristics			
% unemployed above state median			0.06
			(0.08)
% maternal labor force above state median			-0.06
			(0.07)
% Black, non-Hispanic above state median			-0.05
			(0.07)
% Hispanic above state median			0.04
			(0.08)
Median household income above state median			0.24**
			(0.07)
No. of children under 5 above state median			0.11
			(0.10)
Constant	2.20***	2.03***	1.69***
	(0.25)	(0.24)	(0.26)
Observations	481	481	481
Adjusted R-squared	0.357	0.359	0.395

Table D5. Regression analyses for main effects of PD variables predicting 2016-17 CLASS scores, adjusting for program and parish covariates

	Model 1	Model 2	Model 3A
MMCI	0.12*	0.12*	0.15**
	(0.05)	(0.05)	(0.05)
MTP	-0.18	-0.15	-0.18†
	(0.11)	(0.10)	(0.09)
MyTeachStone	0.05	0.06	0.05
	(0.06)	(0.06)	(0.06)
CLASS PreK Reliability Training	0.21**	0.21**	0.17*
	(0.07)	(0.07)	(0.07)
CLASS Toddler Reliability Training	-0.05	-0.06	-0.03
	(0.08)	(0.08)	(0.08)
GOLD Training	-0.09	-0.08	-0.06
Program Characteristics	(0.11)	(0.12)	(0.10)
2015-16 CLASS	0.55***	0.54***	0.55***
	(0.05)	(0.05)	(0.05)
1 = Head Start		-0.03	-0.00
		(0.11)	(0.09)
% BA+		0.00	0.00
		(0.00)	(0.00)
1 = two classrooms		0.20†	0.28*
		(0.10)	(0.11)
1 = three classrooms		0.23*	0.26*
Parish Characteristics		(0.11)	(0.13)
% unemployed above state median			0.07
			(0.08)
% maternal labor force above state median			-0.00
			(0.08)
% Black, non-Hispanic above state median			-0.06
			(0.07)
% Hispanic above state median			0.04
			(0.07)
Median household income above state median			0.23**
			(0.07)
No. of children under 5 above state median			0.13
			(0.09)
Constant	2.08***	1.90***	1.50***
	(0.25)	(0.26)	(0.25)
Observations	481	481	481
Adjusted R-squared	0.365	0.366	0.405

Table D6. Regression analyses for main and interactive effects of significant PD variables predicting 2016-17 CLASS scores, adjusting for program and parish covariates

	Model 3A	Model 3B	Model 3C
Main Effects			
MMCI	0.15**	0.13**	0.44**
	(0.05)	(0.04)	(0.14)
MTP	-0.18†	-0.19*	-0.16
	(0.09)	(0.09)	(0.24)
CLASS PreK Reliability Training	0.17*	0.17*	0.24**
Interaction Effects	(0.07)	(0.06)	(0.08)
MMCI # MTP			-0.43
			(0.39)
MMCI # CLASS			-0.31†
			(0.16)
MTP # CLASS			0.20
			(0.36)
MMCI # MTP # CLASS			0.12
Program Characteristics			(0.47)
2015-16 CLASS	0.55***	0.55***	0.55***
	(0.05)	(0.05)	(0.04)
1 = Head Start	-0.00	-0.01	0.00
	(0.09)	(0.08)	(0.08)
% BA+	0.00	0.00†	0.00
	(0.00)	(0.00)	(0.00)
1 = two classrooms	0.28*	0.25*	0.24†
	(0.11)	(0.11)	(0.13)
1 = three classrooms	0.26*	0.25†	0.24
Parish Characteristics	(0.13)	(0.13)	(0.15)
% unemployed above state median	0.07	0.07	0.09
	(0.08)	(0.07)	(0.07)
% maternal labor force above state median	-0.00	-0.02	0.01
	(0.08)	(0.07)	(0.07)
% Black, non-Hispanic above state median	-0.06	-0.06	-0.06
	(0.07)	(0.07)	(0.07)
% Hispanic above state median	0.04	0.04	0.02
	(0.07)	(0.07)	(0.07)
Median household income above state median	0.23**	0.24**	0.26***
	(0.07)	(0.07)	(0.07)
No. of children under 5 above state median	0.13	0.12	0.11
	(0.09)	(0.09)	(0.09)
Constant	1.50***	1.49***	1.41***
	(0.25)	(0.23)	(0.23)

Appendix E. Equity Estimations

To assess equity, I build off Equation (2) and fit the regression equation on a split sample. For example, to assess equity on median household income, I split the sample to greater than and less than the median income, and fit Equation (2) in each sample. This allows me to estimate the coefficient in each sample and compare the relative magnitudes. I prefer this approach relative to an interactive model (i.e., interacting the policy variable with the demographic variable): although it is simple math to translate coefficients from the interactive model into those presented her, it is nevertheless easier to present the magnitude of the coefficient for each sample.

To understand the extent to which a policy option may be more effective for one group, I conduct seemingly unrelated estimations using Stata 13.1's `suest` command, which accounts for the fact that estimates obtained from two regressions run separately may be correlated with each other. This also allows me to obtain a test statistic to compare coefficients across models. The determination of equity is based on the direction of β_5 and which subgroup the interaction effect favors (see Table 4).

Table E1. Regression analyses for heterogeneity in main effects of significant PD variables predicting 2016-17 CLASS scores, adjusting for program and parish covariates

	% Black, Non-Hispanic			**% Hispanic**			**Median Income**			**QRIS Rating**		
	>50% Black	< %50 Black	Sig?*	>50% Hispanic	<50% Hispanic	Sig?*	< Median Income	> Median Income	Sig?*	Level 1 & 2	Level 3 & 4	Sig?*
Main Effects												
MMCI	0.21	0.35**	ns	0.17†	0.66***	*	0.44*	0.20*	ns	0.28†	0.54**	ns
	(0.14)	(0.13)		(0.10)	(0.20)		(0.22)	(0.10)		(0.17)	(0.19)	
CLASS PreK Reliability Training	0.09†	0.44***	**	0.22*	0.28†	ns	0.24*	0.18†	ns	0.27*	0.27*	ns
	(0.05)	(0.12)		(0.08)	(0.17)		(0.12)	(0.10)		(0.11)	(0.12)	
Interaction Effects												
MMCI # CLASS	-0.10	-0.39**		-0.14	-0.46*		-0.29	-0.15		-0.21	-0.43*	
	(0.15)	(0.15)		(0.12)	(0.23)		(0.22)	(0.13)		(0.17)	(0.20)	
*Program Characteristics***	Y	Y		Y	Y		Y	Y		Y	Y	
*Parish Characteristics***	Y	Y		Y	Y		Y	Y		Y	Y	
Constant	1.68***	-1.43***		1.71***	-1.67***		-1.64***	1.75***		3.42***	4.20***	
	(0.20)	(0.12)		(0.32)	(0.11)		(0.11)	(0.30)		(0.21)	(0.19)	
Observations	289	217		364	142		179	327		248	233	

Notes. Outcome variable in all regressions is 2016-17 CLASS scores.
* Sig. denotes whether the pair of coefficients at left are statistically significantly different from each other. Symbols next to coefficients denote whether that individual coefficient is statistically significantly different from zero. Significance: † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$
** All program and parish covariates are included in each model, EXCEPT the variable on which the sample is split in half. That is, estimations for median household income omit the median household income variable on the righthand side.

Table E2. Regression analyses for heterogeneity in main effects of significant CDA variables predicting 2016-17 CLASS scores, adjusting for program and parish covariates

	% Black, Non-Hispanic			% Hispanic			Median Income			QRIS Rating		
	>50% Black	< %50 Black	Sig?*	>50% Hispanic	<50% Hispanic	Sig?*	< Median Income	> Median Income	Sig?*	Level 1 & 2	Level 3 & 4	Sig?*
Main Effects												
50% or more with CDA	0.08	0.25***	ns	0.16**	0.14	ns	0.10	0.19*	ns	0.26**	0.13	ns
	(0.11)	(0.07)		(0.05)	(0.20)		(0.12)	(0.08)		(0.08)	(0.14)	
50% or more with LA cert.	-0.01	0.12†	ns	0.11	-0.07	ns	-0.06	0.08	ns	0.03	0.13†	ns
	(0.08)	(0.07)		(0.07)	(0.11)		(0.08)	(0.07)		(0.17)	(0.07)	
Interaction Effects												
CDA # LA cert.	0.16	-0.20		-0.06	0.06		-0.01	-0.04		0.30	-0.11	
	(0.17)	(0.15)		(0.10)	(0.32)		(0.18)	(0.14)		(0.28)	(0.18)	
*Program Characteristics***	Y	Y		Y	Y		Y	Y		Y	Y	
*Parish Characteristics***	Y	Y		Y	Y		Y	Y		Y	Y	
Constant	1.74***	2.06***		1.89***	2.21***		1.59***	2.04***		3.61***	4.42***	
	(0.27)	(0.42)		(0.35)	(0.39)		(0.27)	(0.36)		(0.23)	(0.20)	
Observations	270	211		345	136		171	310		248	233	

Notes. Outcome variable in all regressions is 2016-17 CLASS scores.
* Sig. denotes whether the pair of coefficients at left are statistically significantly different from each other. Symbols next to coefficients denote whether that individual coefficient is statistically significantly different from zero. Significance: † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$
** All program and parish covariates are included in each model, EXCEPT the variable on which the sample is split in half. That is, estimations for median household income omit the median household income variable on the righthand side.

Page intentionally left blank