The effect of state quality standards on program quality in Head Start

November 2019

Kathryn E. Gonzalez

Harvard Graduate School of Education

Abstract:

In recent decades, state Quality Rating and Improvement Systems (QRIS) have expanded as a means of assessing and promoting quality in early care and education programs. As QRIS have matured, state and federal policy efforts have focused on incorporating programs across multiple sectors of state early care and education systems, including federally-funded Head Start programs. Using federal administrative data on Head Start programs, I examine how the adoption of QRIS has affected aspects of structural quality, process quality, and teacher turnover in Head Start. Overall, I find that QRIS adoption did not impact several features of structural or process quality in Head Start programs, but did lead to an increase in the annual teacher turnover. Additionally, I find some evidence that impacts on program quality varied based on the approaches used by states to integrate Head Start into the QRIS. Findings point to the potential limitations of existing QRIS as a tool to broadly improve the quality of care in Head Start, and provide an example of potentially unintended consequences of state-level accountability policies for programs that operate at the intersection of state and federal policy contexts.

In recent decades, the provision of high-quality early care and education (ECE) has been a major policy goal at federal, state, and local levels (Barnett, 2003; Barnett & Friedman-Krauss, 2016; Cascio & Schazenbach, 2013; Currie & Neidell, 2007). Yet research from the past decade suggests that the quality of many ECE programs remains persistently low, and that many children do not have access to high-quality, developmentally appropriate early learning environments (Barnett et al., 2015; Bassok, Fitzpatrick, Greenberg & Loeb, 2016; Early et al., 2005). Given the fragmented structure of state ECE systems, including private and publicly-funded programs in center-, school- and home-based settings, relatively few policies have attempted to broadly improve the quality of care. The emergence of state Quality Rating and Improvement Systems (QRIS) represents an effort to improve program quality across ECE sectors. Nearly all states have adopted a QRIS in order to standardize quality and incentivize quality improvement in ECE programs. These systems allow states to define comprehensive quality standards, deliver quality improvement supports, and transparent information on program quality to help families choose high-quality providers (Tout et al., 2010).

Federally funded Head Start programs are a distinct sector of ECE systems that provide a range of education, health, and other services to low-income children and families. Unlike many other sectors of ECE systems, Head Start programs must meet relatively stringent federal quality standards; the Head Start Program Performance Standards (HSPPS; Administration for Children & Families, 2016). At the same time, Head Start programs currently participate in nearly all state QRIS systems (Tout et al., 2010). Additionally, there is an emphasis on increasing Head Start participation in state QRIS at the federal level, including recent revisions to the HSPPS requiring QRIS participation for many programs by 2019 (Administration for Children & Families, 2016).

Head Start programs are incorporated into states' QRIS under the assumption that QRIS participation will support these programs' quality improvement efforts. However, it is unclear whether this assumption is likely to hold given the potential for state QRIS standards to be redundant with federal Head Start requirements. Additionally, the extent to which QRIS generate quality improvements among Head Start programs likely varies based on how states design their QRIS and how states choose to incorporate Head Start programs into their systems. Yet there is little understanding of how state QRIS adoption has affected program quality among Head Start programs, and the extent to which impacts vary across states.

The present study contributes to the understanding of the impacts of QRIS on state ECE systems by examining how states' adoption of QRIS has impacted multiple aspects of program quality in Head Start. In doing so, I provide evidence of whether the demonstrated effects of QRIS adoption on program quality in the center-based sector more broadly (e.g., Bassok, Dee & Latham, 2019; Herbst, 2018) generalize to specific programs such as Head Start that face external quality requirements. First, I conduct a difference-in-differences analysis, utilizing variation in the timing of QRIS adoption across states, to examine the impact of state QRIS adoption on structural quality, process quality, and teacher turnover in Head Start. Second, I conduct moderation analyses to explore whether the impact of QRIS adoption differed across states based on their approach toward incorporating Head Start into the QRIS, focusing on whether states use a different approach to rate Head Start programs as compared with other center-based programs. Third, I examine whether there is evidence for various demand- and supply-side mechanisms through which QRIS adoption is hypothesized to promote quality improvement (e.g., changes in enrollment among high- and low-quality programs and the provision of supply-side supports and incentives). Finally, in light of unexpected patterns of

impacts on Head Start teacher turnover, I explore the possibility of general equilibrium effects of QRIS adoption via changes in compensation among non-Head Start programs.

Overall, I find little evidence that QRIS adoption has broadly improved program quality in Head Start. Results indicate that QRIS adoption had little overall impact on teacher qualifications or the quality of teacher-child interactions in Head Start programs. However, I also find some evidence that design of states' QRIS moderated the impacts of QRIS adoption. Specifically, I observe larger increases in Head Start teacher qualifications among states requiring Head Start programs to comply with the same set of rating procedures as other center-based programs, relative to impacts in states that used rating procedures intended to accelerate or streamline Head Start participation in QRIS. This suggests that the lack of overall impacts on program quality in Head Start may be due, in part, to state QRIS not incentivizing Head Start programs to meet quality requirements beyond the programs' external federal standards.

In contrast, I find evidence that QRIS adoption increased the annual rate of teacher turnover in Head Start by approximately one percentage point. Exploratory analyses indicate the impacts on turnover were larger in states with more parity in compensation between Head Start and other ECE programs. This points to the possibility that QRIS impacts on Head Start turnover may to be due, in part, to general equilibrium effects results from increased compensation in non-Head Start ECE programs, although I cannot rule out other potential explanations for this pattern of effects (e.g., that teachers who do not meet QRIS standards are leaving ECE labor market). Overall, findings suggest the need for caution when considering QRIS as a policy tool to broadly improve quality in Head Start. These findings also indicate the potentially unintended impacts of state accountability policy on sectors of ECE systems that operate at the intersection of state and federal policy contexts.

Background

Overview of Quality Rating and Improvement Systems

QRIS are multi-faceted systems that include a combination of demand- and supply-side interventions. Although the specific components of QRIS vary across states, Figure 1 presents a general framework of the mechanisms through which QRIS are hypothesized to promote quality. The core components of QRIS are quality ratings for providers, which are based on multiple quality indicators selected by the state (Mitchell, 2009). States' QRIS ratings are often include measures of *structural quality* (e.g., class sizes and teacher qualifications) as well as measures of *process quality* (e.g., classroom observations of teacher-child interaction quality; Connors & Morris, 2015; NCECQA, 2017a). These quality indicators are aggregated into program ratings – often a "star rating" or "tier" of quality – which are provided to parents, typically through online resources such as state websites or through provider outreach (Tout et al., 2010).

Empirical evidence suggests that parents often do not have accurate information about the quality of their children's early education programs (Bassok, Markowitz, Player & Zagardo, 2017). The theory of QRIS is that the availability of these ratings will help parents choose high-quality ECE providers, increasing the demand for high-quality care. As a result, low-quality providers will make efforts to obtain higher quality ratings by investing in improvement efforts in areas aligned with the composition of QRIS ratings (e.g., by hiring more teachers with advanced degrees or reducing class sizes) or else exit the market (Cannon, Zellman, Karoly & Schwartz, 2017). In addition, most QRIS also incorporate supply-side intervention components to encourage quality improvement. Typically, this includes a combination of financial incentives tied to programs' participation in and progress through the QRIS, as well as the provision of technical assistance (TA; Mitchell, 2012). Financial incentives often include a combination of

quality awards, which provide funding to programs that reach the higher levels of the QRIS, and improvement grants, which provide funding to help programs advance to the higher levels of the QRIS. Many states also offer tiered reimbursement tied to QRIS ratings, which provide higher subsidy payments to programs with higher QRIS ratings (NCECQA, 2017b). Technical assistance provided to participating programs includes, for example, assistance with development quality improvement plans (Mitchell, 2012).

The earliest QRIS were adopted by a small number of states in the late 1990s and 2000s. These early QRIS were among the first large-scale efforts to encourage ECE programs to reach levels of quality beyond minimum state licensing standards. While early QRIS focused on the provision of quality ratings in order to inform parents' decision-making, the focus of these systems has since expanded to emphasize supporting quality improvement and developing unified ECE systems (Cannon et al., 2017). More recently, the federal Race to the Top Early Learning Challenge (RTT-ELC) led to a dramatic increase in states' adoption of QRIS. In 2005, fewer than 10 states had adopted a QRIS; as of the beginning of 2017, 42 states, 3 Florida counties, and the District of Columbia had adopted a QRIS. Although rates of provider participation in QRIS vary across states, information collected by the BUILD Initiative, a multi-state partnership created to support states' development of ECE systems, indicates that rates of QRIS participation among center-based providers are generally high (i.e., often above 50 percent; BUILD Initiative, 2017).

Evidence Regarding Quality Improvement in QRIS

Most of the research to date on QRIS has focused on the validity of QRIS ratings in order to determine whether higher-rated programs provide higher-quality experiences for children (Hestenes et al., 2015; Sabol, Hong, Pianta, & Burchinal, 2012; Sabol & Pianta, 2015). A more

recent literature has begun to examine how families and providers respond to states' adoption of QRIS. Research provides some evidence to suggest that the receipt of QRIS quality improvement supports is associated with subsequent quality improvements (e.g., Boller et al., 2015; Yazejian & Iruka, 2015), although much of this evidence is correlational or in the context of pilot QRIS rather than systems that are fully implemented at scale. More recently, Bassok, Dee, and Latham (2019) provide evidence regarding the efficacy of the combination of incentives and supports for quality improvement created through the receipt of different quality ratings. The authors find that in North Carolina, the quasi-random assignment to a lower quality rating led to subsequent improvements in program quality.

However, few studies have examined the overall impact of states' QRIS adoption on program quality. Most recently, Herbst (2018) examines how states' QRIS adoption broadly impacted the supply and quality of labor in ECE programs. The author finds that QRIS adoption increased the supply of child care labor in center-based programs by individuals with at least some college education and increased monthly earnings of child care workers, but also increased employee turnover. However, there is little understanding of whether QRIS adoption broadly affected other aspects of structural quality (e.g., class sizes) and process quality (e.g., the quality of teacher-child interactions). Additionally, state ECE systems include a wide variety of programs that vary widely in term of structure, compensation, and external requirements outside the QRIS (CSCCE, 2017). For example, Head Start programs must meet federal standards that provide guidelines for various aspects of quality as well as ongoing program monitoring (US HHS Administration for Children & Families, 2016). It is not clear whether observed impacts of QRIS on program quality generalize to these programs.

Federal Quality Standards for Head Start

Head Start comprises a unique sector of states ECE systems. As the first and largest federally-funded early education program, Head Start serves roughly 900,000 low-income children and families each year. Several decades of research have highlighted positive impacts of participation in Head Start on school readiness (Puma et al., 2010), health and behavioral outcomes (Carneiro & Ginja, 2014; Ludwig & Miller, 2007), and on longer-term economic well-being (Bauer & Schanzenbach, 2016; Deming, 2009). However, research also documents considerable variation across Head Start programs and classrooms in terms of observed measures of quality as well as program impacts on children's school readiness (Bloom & Weiland, 2015; Sabol, Ross & Frost, 2019; Walters, 2015).

Unlike many other types of ECE programs, Head Start programs must meet relatively stringent federal quality standards, known as the Head Start Program Performance Standards (HSPPS). The HSPPS mandate requirements for programs' operation and governance, such as child eligibility and enrollment, program structure, education services, and health services (Administration for Children and Families, 2016). Revisions of these standards over the past few decades have sought to broadly improve the quality of care in Head Start programs. For example, as a result of the 2007 reauthorization of the HSPPS, 50 percent of Head Start teachers were required to have a Bachelor's degree by 2013 (Administration for Children and Families, 2007). Since 2012, as part of the implementation of the Head Start Designation Renewal System (Head Start DRS), Head Start programs have also been evaluated on aspects of process quality using the Classroom Assessment Scoring System (CLASS; Pianta, La Paro & Hamre, 2008). Grantees that receive low scores must re-compete for funding (Administration for Children and Families, 2016b). Most recently, the 2016 HSPPS reauthorization requires most Head Start programs participate in state QRIS (Administration for Children and Families, 2016a).

Head Start Participation in State QRIS

QRIS participation is typically voluntary (Tout et al., 2010). However, there is some evidence that Head Start participation is typically high. Currently, Head Start programs are eligible to participate in most states' QRIS (BUILD Initiative, 2017). A recent study of seven states that participated in the RTT-ELC estimated that, as of 2016, approximately 90 percent of state-funded prekindergarten and Head Start programs participated in their states' QRIS (Kirby, Herrmann, Wolfendale & Mraz Esposito, 2019). However, the extent to which QRIS adoption can promote quality improvement in Head Start likely depends in part on the extent to which state QRIS standards exceed the federal HSPPS. For example, descriptive evidence suggests that in some states Head Start programs frequently enter QRIS at the highest levels, leaving limited room for improvement (Herrmann et al., 2019; Kirby et al., 2019).

This pattern may be driven in part by states' efforts to incorporate Head Start requirements into state QRIS standards, towards the aim of developing a more unified system of standards across ECE sectors (BUILD Initiative, 2018; Herrmann et al., 2019). Many states have developed crosswalks between QRIS standards and HSPPS requirements; in some states HSPPS requirements meet or exceed requirements for the highest QRIS rating levels (BUILD Initiative, 2018). For example, nearly all QRIS ratings include teacher qualifications (NCECQA, 2017c). However, these standards do not typically exceed the 2007 federal requirement that at least 50 percent of Head Start teachers have Bachelor's degrees[1]. Therefore, it is unlikely that QRIS adoption would have affected this aspect of quality among Head Start programs after the 2007 federal requirement went into effect.

[1] The classifications of QRIS standards are based on a review of quality standards from state websites.

However, in other areas QRIS standards may exceed the HSPPS. For example, many states have requirements around teacher-child interaction quality that exceed the Head Start Program Performance Standards. For example, 21 QRIS currently incorporate CLASS scores as part of programs' QRIS ratings (BUILD Initiative, 2018). Of these, at least 13 states require programs to meet minimum thresholds in the CLASS to obtain certain quality ratings[2]. In most of these states, CLASS thresholds in required to reach the highest levels of the QRIS exceeded both the minimum absolute and relative CLASS score requirements for Head Start programs under the DRS[3]. Therefore, it plausible that QRIS adoption may have impacts on some, but not all, aspects of quality in Head Start programs.

In addition, many state QRIS have also adopted different procedures to rate Head Start programs as compared with other center-based programs. In some states, Head Start must meet the same quality standards as other center-based programs (i.e., Head Start programs follow the *standard rating process*). Other states use alternative rating processes for Head Start programs that participate in the QRIS (i.e., Head Start programs use what I refer to as an *accelerated rating process;* Tout et al., 2010). As of 2017, 25 states reported that their QRIS include an "accelerated" rating process for participating Head Start programs[4]. In general, states use two types of accelerated rating processes. The first approach is the use of "reciprocity" for Head Start programs. Under this approach, Head Start programs in good federal standing automatically meet

[2] This was based on a review of the *Quality Compendium* and state websites. Programs without minimum thresholds either used a point-based system, required programs to conduct CLASS observations but did not set minimum scores, or did not have information about CLASS usage available.

[3] For example, the Head Start DRS requires Head Start grantees to score at least a 4 on emotional support, a 3 on classroom organization, and a 2 on instructional support (Administration for Children and Families, 2016b). For all 13 QRIS where information was available, the CLASS scores needed to obtain the highest QRIS rating exceeded those thresholds, across all three domains. The DRS also programs with CLASS scores in the bottom 10 percent to re-compete for funding; in 2014, this represented a score of 5,7 on emotional support, a 5.3 on classroom organization, and 2.2 on instructional support. In most cases, requirements for the highest QRIS rating exceeded these levels as well.

[4] This was based on a review of the *Quality Compendium* and provider rating guides from state websites.

specific quality indicators aligned with the HSPPS with little or no additional verification. Head Start programs must then go through the full verification process only for QRIS standards not covered by the HSPPS. For example, Oregon and Montana both allow Head Start centers to use federal monitoring data and compliance with federal Head Start standards to meet certain QRIS standards (BUILD Initiative, 2018). The QRIS monitoring burden for Head Start programs is therefore reduced when there is considerable overlap between the two sets of standards.

The second primary approach of accelerating Head Star participation in state QRIS is the use of "alternative pathways" to ratings. In these states, the processes used to evaluate Head Start programs are substantively different (and generally less intensive) than those used for other center-based programs. Often, Head Start programs in these states automatically attain certain QRIS ratings by virtue of being in good federal standing. For example, in Minnesota and Illinois, Head Start programs are automatically eligible to receive high ratings based on the submission of federal Head Start monitoring reports. Other states assign ratings to Head Start programs using an abbreviated set of standards that are less comprehensive than the indicators used to rate other center-based programs (BUILD Initiative, 2018).

There is little empirical evidence regarding whether programs that use an accelerated QRIS rating process would obtain the same ratings had they gone through the full rating process (Kirby et al., 2018). Requiring Head Start programs to use standard rating processes – rather than, for example, granting automatic ratings – could lead to improvements in areas where state QRIS standards exceed HSPPS requirements. However, studies find mixed evidence regarding whether there are differences in observational quality among highly-rated programs used accelerated rather than standard (Karoly et al., 2016; Tout et al., 2016; Roberts et al., 2016).

Therefore, understanding of the implications of states' decisions around how to incorporate Head Start programs into their QRIS in terms of supporting quality improvement remains limited.

Timing of QRIS Adoption Across States

As of 2017, 42 states, 3 Florida counties, and the District of Columbia have a QRIS[5]. Prior to full implementation of a QRIS, many states operated small, pilot systems. In the present study I focus on the impacts of full QRIS adoption when the systems were first implemented statewide (or county-wide). I collected QRIS start dates from the *Quality Compendium*, a web-based database of all state QRIS that is gathered and maintained by the BUILD Initiative, as well as a previous study of QRIS impacts (Herbst, 2018). I then cross-referenced the dates of QRIS adoption with other sources including reports and state websites [6].

As shown in Figure 2, there is considerable variation in the timing of QRIS adoption across states. In 2000, three states had adopted a QRIS. This number increased to 22 in 2010 and was over 40 by 2016. This acceleration was due in part to the federal Race to the Top Early Learning Challenge (RTT-ELC), which provided funding to states to improve young children's access to high-quality early learning programs. States applying to the RTT-ELC were required to demonstrate the use or development of a tiered rating system (i.e., a QRIS) in order to receive funding. Between 2012 and 2014, 20 states were named as recipients RTT-ELC grants

[5] I include Florida and California in this number. California operates a statewide QRIS on a county-by-county basis. Three Florida counties operate a QRIS: Duval County, Miami-Dade County, and Palm Beach County. For simplicity, I refer to the District of Columbia as a state for the remainder of the paper.

[6] The BUILD Initiative *Quality Compendium* is available at: https://qualitycompendium.org/). Notably, the year of QRIS adoption used in the present study differs from Herbst (2018) for two states. First, I classified Oregon as having adopted a QRIS in 2014. This was the year in which the state-wide field test of the Oregon QRIS began. This field test is treated as a fully implemented QIRS for the purpose of this analysis due to scale of the field test (including over 34,000 children enrolled in participating providers as of January 2015) in combination with the fact that QRIS ratings have been made available to parents on the state's website. Second, I classified Texas as having adopted a QRIS in 2015. This was the year in which the state revised QRIS guidelines. In 1991, Texas issued first provider certification standards for child care providers through the Texas Rising Star (TRS) program. The TRS program was then used as the basis for the statewide Texas QRIS as part of a revision of the TRS program; these revisions were approved in January 2015.

(Administration for Children and Families, 2017); roughly half of those states adopted a statewide QRIS between 2011 and 2014.

Only four states have limited or no Head Start QRIS participation: Alaska, North Dakota, Texas, and Utah[7]. However, Head Start programs in these states may nevertheless have been affected by their state's adoption of QRIS. For example, QRIS adoption may have affected teachers' decision about whether to teach in Head Start or in other ECE sectors. Alternatively, Head Start programs may have been encouraged to meet state QRIS standards due to increased parental demand for higher-quality care. Therefore, I treat all states and counties with a QRIS as treated states in the present analysis. As a robustness check, I also replicate the primary analyses excluding states where Head Start programs are not eligible to participate in QRIS.

Methods

Data and Sample

Data for this study primarily came from two sources maintained by the U.S. Department of Health and Human Services, Administration for Children and Families (ACF). This includes Head Start Program Information Reports (2003-2018)[8] and CLASS Reports (2012-2018). Supplementary data were collected from a range of state and federal sources.

Head Start Program Information Reports (Head Start PIR; 2003-2018). Head Start PIRs are based on annual reports submitted by Head Start grantees and contain panel data on all

[7] North Dakota is currently involved in discussions regarding bringing Head Start programs into the system (BUILD Initiative, 2017). In Texas, only programs with an agreement to serve children receiving subsidies can participate in the QRIS. As a result, there is limited Head Start participation in the Texas QRIS (BULID Initiative, 2017). This was confirmed by an examination of the lists of providers currently participating in the QRIS provided by the state (https://texasrisingstar.org/parents/find-a-trs-provider/), which included few Head Start programs. In Utah, only licensed programs are eligible to participate; Head Start participation is therefore limited (BUILD Initiative, 2017). The license-exempt status of most Head Start programs was confirmed by an examination of the lists of providers operating in Utah (https://careaboutchildcare.utah.gov/parent/search.aspx).

[8] I focus on 2003 and later due to the fact that additional information about states' early education and care programs are available starting in 2003.

Head Start programs. Grantees receive federal funding to operate one or more Head Start centers or classrooms. In some cases, grantees who received federal Head Start funding subsequently provided the funding to one or more delegate agencies (i.e., programs); the unit of observation in the PIRs is at the grantee-delegate agency-year level (i.e., program-year level). The data include yearly program-level information on teacher, staff, and enrollment information, supporting comparisons of program quality across states which may differ in the structure of their ECE systems. The present analysis focuses on all Head Start programs that enrolled children in center-based settings (including American Indian and Alaska Native Head Start [AIAN HS]) and Migrant and Seasonal Head Start programs [MS HS]) from 2003 to 2018. I exclude Early Head Start programs as well as programs that did not report operating at least one classroom, employing at least one preschool classroom teacher, and enrolling at least one child in a center-based setting[9]. The final sample includes 28,571 program-year observations.

Head Start CLASS Reports (2012-2018). Since 2012, a nationally representative samples of Head Start grantees have been evaluated annually using the CLASS as part of the Head Start DRS. The CLASS is one of the most widely used measures of process quality in ECE settings and is described in detail below. In each year between 2012 and 2018 roughly 300 to 400 grantees were observed using the CLASS; few grantees were observed more than once[10]. Unlike the Head Start PIRs, CLASS reports include information at the grantee level rather than the program level. For the present analysis, CLASS scores for 2,021 grantee-year observations were

[9] Head Start grantees that reported having zero classroom teachers typically include grantees where all reported enrollment was in home based or family child care options.

[10] As the present analysis relies on state-level variation in the key predictor of interest (QRIS adoption), a concern is that the relatively small number of grantees observed in some states could lead to noise in the impact estimates. Additionally, measures of quality among smaller states with fewer grantees would likely contain more error. Although this may lead to imprecision, measurement error in the outcome variable should not bias impact estimates.

web-scraped from the ACF website. Of these, 15 could not be matched to grantees the Head Start PIRs. This yielded a final sample of 2,006 grantee-year observations.

Additional sources of information on state QRIS. Information on the components of state QRIS was compiled primarily from the *Quality Compendium* maintained by the BUILD Initiative. The *Quality Compendium* includes detailed information on the components of each state's QRIS, including the use of accelerated rating processes for certain types of programs, the types of financial incentives and TA provided to participating programs, and the specific quality indicators used to determine QRIS ratings. Additional information was also collected from other online and published resources to confirm the information in the *Quality Compendium*.

Key Measures

Structural quality and teacher turnover outcomes. Outcomes of interest generated from the Head Start PIRs include two measures of structural quality that are typically incorporated into states' QRIS ratings, teacher qualifications and group sizes, as well as annual rates of teacher turnover. Teacher qualifications were measured by the share of classroom teachers (excluding assistant teachers) in each program that had a Bachelor's degree or higher[11]. Group sizes were calculated by the ratio of total child enrollment in center-based settings and the total number of classrooms. The rate of annual teacher turnover was calculated by the number of classroom teachers (excluding assistant teachers) who were reported to have left during the program year (including during the academic year and summer months) divided by the total number of classroom teachers. In addition to information on overall teacher turnover, the PIR

[11] Migrant and Seasonal Head Start programs operate both preschool and infant and toddler classrooms. Prior to 2011, teacher qualification information was not reported separately for teachers in preschool and infant and toddler classrooms. Therefore, in all years, measures of teacher qualifications for Migrant and Seasonal Head Start programs include teachers in both preschool and infant and toddler classrooms. In all years, measures of teacher qualifications for Head Start and American Indian Alaska Native Head Start programs include only preschool teachers.

also report the number of teachers who left their program for each of the following: higher compensation or benefits in the same field, change in job field, and other reason[12]. In addition to overall annual teacher turnover, I also calculate rates of annual turnover due to each reason.

As of 2018, programs reported that the share of teachers with a Bachelor's degree or above in their center(s) was 69 percent, on average, up from 36 percent in 2003. This growth aligns with other reported statistics regarding Head Start teacher qualifications (e.g., Aikens et al., 2016). The average group size remained constant between 2003 and 2018, while rates of annual teacher turnover increased from 15 percent to 19 (see Table 1).

Process quality outcomes. Grantee-level CLASS from the Head Start CLASS reports comprise the process quality outcomes of interest. The CLASS measures the quality of teacher-child interactions in three domains. *Emotional support* captures the extent to which teachers provide a positive and emotionally supportive classroom. *Classroom organization* captures teachers' management of the classroom and use of strategies to encourage learning. *Instructional support* captures teachers' provision of feedback and stimulation of language in the classroom (Pianta, La Paro & Hamre, 2008). Outcomes of interest include the three CLASS domain scores, which are scored on a scale of one to seven. CLASS scores are standardized with respect to the sample mean and standard deviation for analysis.

According to the ACF, 2,411 grantees received CLASS observations over this time period[13]. Therefore, the analysis sample for the present study represents 83 percent of CLASS observations conducted over this time period. The percent of grantees in the present sample relative to the total number of conducted observations in each year ranged from 69 percent

[12] Of teachers who left Head Start programs, approximately 30 percent left because of higher compensation/benefits in another program, 18 percent left to change fields, and 53 percent left for other reasons.
[13] This information was collected from the national statistics of CLASS scores published by the Administration for Children and Families in each year. Reports are available on the ACF website: https://eclkc.ohs.acf.hhs.gov.

(2014) to 94 percent (2015). Missing CLASS scores may be due to grantees that left the system between when the CLASS observation was conducted and when the web-scraping was conducted in 2017 and 2018. Other CLASS scores may not have been available on the ACF website. A potential concern is that the grantees with missing CLASS reports may have been disproportionally low-quality. However, the average CLASS scores across the three domains align closely with the yearly national averages reported by the ACF (see Appendix Table A1).

QRIS rating processes for Head Start. I classified each state's QRIS based on whether states streamline Head Start programs' participation in the QRIS. States which utilized either reciprocity or alternative rating pathways for Head Start centers were classified as "accelerated rating" states. States which did not use either of these pathways were classified as "standard rating" states. In addition, some states (particularly early-adopting states) adopted a QRIS that initially required Head Start programs to follow a standard rating process, but subsequently implemented accelerated rating processes. Therefore, I classify states along this dimension based on their use of accelerated rating processes in a given year[14]. The number of states using each type of rating processes increased steadily between 2002 and 2018 (see Appendix Figure B1).

QRIS supply-side components. I also classified each state's QRIS based on the supply-side interventions incorporated into the QRIS. In general, these include a combination of financial incentives, typically given to providers reaching or moving towards certain levels of quality, as well as technical assistance (TA) to support programs' quality improvement efforts. I classified each state's QRIS based on whether it incorporates the following financial incentives:

[14] This information was determined based on documentation from state websites – for example, policy briefs describing changes to the QRIS rating procedures – and state entries in the *Quality Compendium* (including archived copies of the database from 2014 through 2017) that ask states to describe the use of accelerated/automatic rating processes as well as major revisions to the QRIS. However, this information was not available for most states; in these cases, I assume that the state had an accelerated rating process in all years.

quality improvement grants, quality awards and bonuses, and teacher scholarships. I also classified QRIS based on whether states provide the following types of TA: mentoring, coaching, consultation, professional development (PD) advising, and peer-to-peer TA.

Teacher compensation. In order to conduct supplemental analyses of the moderating role of broader ECE labor market conditions, I also collected state-level information on the average hourly wages of child care workers and preschool teachers from 2004 through 2018 (adjusted for inflation to 2018 dollars) from the Bureau of Labor Statistics Occupational Employment Statistics (OES) Survey. Additionally, I collected state-year level information on the average hourly wages of Head Start teachers from the Head Start PIRs.

Time-varying state characteristics. I include a series of time-varying state covariates in all analyses as controls. This includes measures of state-level economic conditions, including the annual unemployment rate, female employment rate, and per-capita income (obtained from the Bureau of Labor Statistics). I also included a series of time-varying state characteristics that represent other aspects of states' ECE landscape, including federal Head Start funding to the state (obtained from the ACF), as well as the percent of four-year-olds enrolled in state-funded prekindergarten programs and state spending per child enrolled in state-funded prekindergarten (obtained from annual yearbooks published by the National Institute for Early Education Research [NIEER]). All dollar values were adjusted for inflation to 2018 dollars.

Time-varying program characteristics. I also include a series of time-varying grantee covariates in all analyses as controls, based on information reported in the PIRs. This includes the ages of enrolled children (i.e., percent three years old, percent four years old, and percent five years old or older), the race/ethnicity of enrolled children (i.e., percent black, percent white, percent Hispanic, and percent American Indian/Alaska Native), percent of enrolled children

whose home language was English, and program type (i.e., Head Start, AIAN HS, MS HS). For analyses of CLASS scores, which were collected at the grantee level, I calculated the average level of each covariate across programs operated by the grantee.

Analytic Approach

Overall impacts of QRIS adoption on program quality in Head Start. I use a differences-in-differences approach to examine the impact of QRIS adoption on program quality in Head Start, utilizing variation in the timing of adoption of QRIS across states. This approach controls for all time-invariant state characteristics through the inclusion of fixed effects for the 43 states and the 3 counties with a QRIS. I treat the three counties as separate states in the present analysis, and for simplicity I refer to these as state fixed effects. This approach also controls for any contemporaneous nationwide policy changes through the inclusion of year fixed effects (e.g., changes to the federal HSPPS) that affect all Head Start programs nationwide. In order to identify causal impact of QRIS adoption, this difference-in-differences approach relies on the assumption of parallel trends. That is, the change in program quality over time that occurred among states without QRIS is the same change that would have occurred in states that adopted a QRIS in the absence of these systems. I conduct falsification tests for my primary analyses to examine whether the assumption of parallel trends is reasonable.

In the first set of analyses, I compare within-state changes over time in structural quality, process quality, and teacher turnover between Head Start programs in states that adopted QRIS systems relative to states that did not. The primary analysis takes the following form:

$$(1)\ Y_{ist} = \alpha + \beta QRIS_{st} + \delta X_{ist} + \rho W_{st} + \mu_s + \lambda_t + \epsilon_{ist}$$

where Y_{ist} represents a program quality outcome in Head Start for program *i* in state *s* and year *t*, including aspect of structural quality, process quality, and teacher turnover. The variable $QRIS_{st}$

is the predictor of interest, and is an indicator for whether state *s* had a QRIS in year *t.* The model also includes a vector of time-varying program characteristics, X_{ist}, and a vector of time-varying state-level characteristics, W_{st}. I also include state fixed effects (μ_s) and year fixed effects (λ_t). Therefore, β is the coefficient of interest and represents the overall impact of QRIS adoption.

I also estimate event study models in order to examine the dynamics of QRIS impacts. For example, it may be the case that impacts of QRIS adoption increase over time as more providers join the system and have time to make quality improvement efforts. I therefore estimate a version of equation (1) in which the QRIS indicator is replaced with a series of binary variables indicating leads and lags of QRIS adoption ranging from 10 or more years before QRIS adoption to 10 or more years after QRIS adoption. As process quality information is only available in the limited window from 2012 to 2018, these analyses include only the structural quality and teacher turnover outcomes. I also examine whether impacts of QRIS adoption varied among early-adopting versus late-adopting states. For example, changes in federal Head Start policy (e.g., revisions to the HSPPS) may have affected the extent to which QRIS could support quality improvement. Therefore, I also estimate several versions equation (1) comparing states that adopted a QRIS in a given year with states that adopted a QRIS several years later or states that never adopted a QRIS. I conducted separate analyses for each year of QRIS adoption.

Impact variation based on use of accelerated rating processes for Head Start. It is important to note that the estimates described above reflects the *average* impacts of QRIS adoption. As states' QRIS vary widely in terms of structure and design, these impacts may mask considerable impact heterogeneity across states. In a second set of analyses I examine whether the impacts of QRIS adoption vary based on whether states use standard rather than accelerated rating processes for Head Start programs. I compare within-state changes over time in quality

outcomes between Head Start programs in states with accelerated rather than standard rating processes. This moderation analysis takes the following form:

$$(2)\ Y_{ist} = \alpha + \theta QRISAccelerated_{st} + \beta QRIS_{st} + \delta X_{ist} + \rho W_{st} + \mu_s + \lambda_t + \epsilon_{ist}$$

Where $QRISAccelerated_{st}$ is an indicator for whether state *s* had a QRIS with an accelerated rating process (i.e., reciprocity or alternative pathways; presence of specific financial incentives) in year *t*; all other variables are defined as above. Therefore, θ is the coefficient of interest, and represents the marginal difference in impacts of QRIS with accelerated rating processes for Head Start programs relative to impacts of QRIS with standard rating processes.

Exploring potential mechanisms: Demand-side QRIS factors. I also examine whether there is evidence to support the hypothesized demand-side mechanisms of QRIS. The theory of QRIS predicts that the availability of ratings should lead to an increase in enrollment among high-quality Head Start programs relative to low-quality Head Start programs as parents choose to enroll their children in high-quality programs and avoid lower-quality programs. I therefore examine whether enrollment in higher-quality Head Start programs increased relative to enrollment in lower-quality programs following QRIS adoption. First, I restrict the sample to programs that were observed both before and after the adoption of QRIS in their state. Second, I classify programs as or high-quality or low-quality based on the programs' quality features in the years prior to QRIS adoption. I consider three quality metrics: teacher qualifications (i.e., percent of teachers with a Bachelor's degree or above), group sizes, and teacher turnover. For a given quality measure, I create an indicator for whether the program was above or below the median relative to other programs in the same state and year. For teacher qualifications, programs were classified as high-quality if they were above the state median in the majority of pre-QRIS years. For group sizes and teacher turnover rates, programs were classified as high-quality if they were

below the state median in the majority of pre-QRIS years. Finally, I estimate a difference-in-differences model that compares within-program changes in higher-quality versus lower-quality programs in enrollment before and after QRIS adoption. Formally, I estimate a version of equation (2) that includes an interaction between the QRIS indicator and an indicator for whether a program was high quality, and replace state fixed effects with program fixed effects.

Next, I examine whether impacts of QRIS adoption on program quality and teacher turnover in Head Start differed based on a proxy for the frequency with which families used QRIS ratings to make child care decisions. Specifically, I collected a measure of internet search popularity for each state's QRIS using *Google Trends*[15]. This information provides a measure of how the frequency of internet searches for specific terms changed over time. For each state, I collected information on the frequency of searches for that state's QRIS, using the name of the QRIS (e.g., Arkansas's "Better Beginnings" and Arizona's "Quality First"). I collected monthly search frequency information from 2004 to 2018 for 30 states; this information was aggregated to the state-by-year level by taking the average search popularity across all months of the school year (defined as September to August)[16]. I then estimated a model analogous to equation (2) including the annual measures of search popularity as a time-varying moderator. If providers respond to increases in families demand for high-quality, impacts of QRIS adoption should be larger in years where more internet searches were conducted for the QRIS.

Exploring potential mechanisms: Supply-side QRIS factors. I also examine whether there is evidence to support the hypothesized supply-side mechanisms of QRIS. Specifically, I examine whether states' provision supply-side supports and incentives is associated with larger

[15] Google Trends available here: https://trends.google.com/trends/?geo=US.
[16] The frequency of searches for a state's QRIS increased sharply following the states' adoption; this pattern is more pronounced when states that had a pilot QRIS are excluded (see Appendix Figure C1).

impacts of QRIS adoption. Formally, I estimate a version of equation (2) that includes whether a state's QRIS offered a particular type of supply-side intervention as a moderator of QRIS impacts. I consider the following: technical assistance (including mentoring, coaching, consultation, PD advisor, and peer-to-peer TA) and financial incentives (including quality improvement grants, quality awards and bonuses, and scholarships).

Exploring potential general equilibrium effects: ECE compensation and Head Start teacher turnover. Finally, I examine whether effects of QRIS adoption on other sectors of state ECE systems might impact teacher turnover in Head Start programs, i.e., that there might be general equilibrium effects of QRIS. Head Start programs comprise only a fraction of states' ECE systems (CSCCE, 2017), and non-Head Start providers comprise the majority of providers in most states' QRIS (BUILD Initiative, 2017). Furthermore, impacts of QRIS adoption on non-Head Start sectors of state ECE systems may have implications for Head Start programs. In particular, evidence suggests that QRIS adoption has broadly increased compensation for ECE workers (Herbst, 2018)[17]. However, it is unlikely that similar increases in compensation occurred in Head Start programs, as compensation levels and requirements for educator qualifications in Head Start programs generally exceed those in other center-based ECE programs (CSCCE, 2017)[18]. Therefore, we would expect QRIS adoption to lead to a relatively large increase in the demand for highly-qualified educators (and a corresponding increase in compensation) among non-Head Start programs, and a smaller response among Head Start programs[19]. If increases in

[17] Herbst (2018) estimates that state QRIS adoption increased the monthly earnings of child care workers by about one percent.
[18] Following the 2007 Reauthorization of the Head Start Program Performance Standards, 50 percent of lead Head Start teachers are required to have Bachelor's degree. In contrast, licensing standards for other center-based ECE programs often require little or no formal educational qualifications (CSCCE, 2017).
[19] Compensation could increase if parental demand for high-quality care increase and parents are willing to pay more for high-quality are (Herbst, 2018). Furthermore, financial incentive components of many states' QRIS (e.g., quality awards and bonuses, increased subsidy reimbursement rates) could lead to increases in compensation.

compensation or working conditions were concentrated in non-Head Start programs, some Head Start teachers may have moved to other center-based ECE programs.

To test whether changes in compensation in other sectors of ECE systems could lead to an increase teacher turnover in Head Start, I examine whether parity between Head Start wages, as compared with wages of child care workers and preschool teachers, moderates the impact of QRIS adoption on Head Start teacher turnover. First, I calculate the difference in hourly wages between Head Start teachers and other ECE educators in each year. I then include this measure of compensation parity as well its interaction with the QRIS indicator to equation (1). In a second specification, I interact the QRIS indicator with the within-state *change* in the average wage gap before and after QRIS adoption[20].

Evidence Regarding the Parallel Trends Assumption

A fundamental assumption underlying the interpretation of the primary difference-in-differences estimates as the causal impact of QRIS adoption is the parallel trends assumption. That is, that the changes in Head Start quality in states without QRIS represent the changes in quality that would have occurred in QRIS states had these systems not be adopted. I provide descriptive evidence regarding this assumption for structural quality and teacher turnover in Figure 3. This figure shows the outcome trends for states that adopted a QRIS in the five years before and after QRIS adoption, centered around the year of QRIS adoption. The left of each figure presents the outcome trends in the years prior to QRIS adoption; the right of the graph presents outcome trends in the years after QRIS adoption. The figure also presents outcome

[20] The change in hourly wage gap was calculated by the following: First, the difference between average child care/preschool and Head Start wages was calculated for each state/county and year. Second, the average wage gap for each state/county was calculated for the years before QRIS adoption (average pre-QRIS wage gap), and the years after QRIS adoption (average post-QRIS wage gap). Third, the change in the hourly wage gap was calculated by the difference between the average pre-QRIS wage gap and the average post-QRIS wage gap.

trends over the same period for states that never adopted a QRIS or adopted a QRIS more than five years after the relevant adoption year. Outcomes trends in pre-QRIS years are generally similar for states with and without a QRIS. Similar figures for process quality outcomes also provide little visual evidence of differences in pre-QRIS trends (see Appendix Figure A1).

To further examine whether the parallel trends assumption is likely to hold in the present study, I also conduct a series of falsification tests to examine whether there is evidence of differences in pre-QRIS trends between states without and without a QRIS. I estimate a difference-in-differences model similar to equation (1) that includes binary indicators for placebo QRIS adoption in the four years (two years for analyses of CLASS outcomes) prior to actual implementation, as well as the primary indicator for states' QRIS adoption. If these placebo treatment indicators are statistically significant or meaningfully large in magnitude, this would suggest that the parallel trends assumption is not satisfied. Results of these falsification tests are presented in Table 2. Consistent with Figure 3, for teacher qualifications, teacher turnover, and process quality the coefficient estimates on the lead QRIS indicators are generally small, and none are statistically significant. In contrast, for group size, coefficients on several placebo treatment indicators are statistically significant. Therefore, there is some concern that estimates of impacts on group sizes may be biased upwards. Due to this potential violation of parallel trends, I do not present results for group size; results are available in Appendix A.

A second key assumption of the difference-in-differences analysis is that unobserved state characteristics are fixed in time. This assumption could be violated if state policies affecting program quality in state ECE systems – or affecting Head Start programs in particular – are enacted concurrently with QRIS. I see little evidence that QRIS adoption was accompanied by changes in state prekindergarten funding, state prekindergarten enrollment, or federal Head Start

funding to the state (see Appendix Table A2). However, another consideration is that many states adopted a QRIS as a result of the federal RTT-ELC program. A potential concern is that these states may have made simultaneous efforts to improve program quality in their ECE systems. However, I find that my results are largely robust (albeit less precisely estimated) when the 20 states participating in RTT-ELC are excluded from the sample, suggesting that changes in state policy driven by RTT-ELC participation are unlikely to be driving the present results.

Results

Overall Impacts of State QRIS Adoption on Program Quality in Head Start

As shown in Table 3, states' adoption of QRIS had no significant overall impacts on the share of teachers with a Bachelor's degree or above in Head Start. Similarly, results provide little evidence to suggest that QRIS adoption generated overall improvements in process quality in Head Start programs. Although point estimates for the impacts of QRIS adoption on emotional support and classroom organization are both positive, both are imprecisely estimated and not statistically significant. The estimated impact on instructional support is negative in sign and similarly not statistically significant.

In contrast, I observe a positive overall impact on teacher turnover, indicating that QRIS adoption increased the annual rate of teacher turnover by 0.9 percentage points ($p < .05$)[21]. This suggests that QRIS adoption increased the annual number of teachers who left Head Start programs. Furthermore, I find evidence that impacts on teacher turnover increased with time

[21] Goodman-Bacon (2018) shows that estimates of treatment impacts can be biased when treatment impacts change over time in a difference-in-differences estimation. The author shows that a difference-in-differences estimate with multiple treated groups at multiple time points is a weighted average of many difference-in-differences groups that compare various groups (e.g., early-adopting, late-adopting, and never-adopting states). In this case, as the results in Figure 4 indicate growing effects of QRIS adoption on teacher turnover over time there are concerns of bias since earlier-adopting states serve as the comparison group for later-adopting states in some of these estimates. This would result in estimates of impacts of QRIS adoption being biased downward for these comparisons. As a result, the present estimate is a possible underestimate of the impact of QRIS on teacher turnover.

since QRIS adoption. Results of estimating an event study model, presented in Figure 4, suggest that impacts on teacher turnover were relatively small in magnitude in the years immediately following QRIS adoption, but increased over time.

One explanation for this result is that QRIS is working as intended, and that teachers who do not meet the QRIS standards are exiting the Head Start market. Although I do not observe concurrent increases in teacher qualifications or process quality, it may be the case that teachers are selecting out of Head Start programs based on other unobserved measures of quality. Alternatively, this finding may indicate unintended consequences for Head Start programs. The increase in turnover could be due to increased accountability pressures that negatively impact program climate or working conditions among Head Start programs. For example, it may be the case that Head Start teachers with Bachelor's degrees are leaving Head Start to teach in state-funded prekindergarten or Kindergarten classrooms where compensation is typically much higher. Alternatively, this finding might be explained by impacts of QRIS on other sectors of state ECE systems (e.g., increased demand for highly qualified teachers or improved compensation) leading teachers to move from Head Start to other ECE programs. I examine this possibility below.

I also confirm that this result is robust to several sensitivity checks. I replicate analyses excluding states which did not adopt a QRIS, excluding the four states which had limited Head Start participation in the QRIS, excluding states which adopted a QRIS before 2003, and excluding states that participated in the RTT-ELC. Results are generally consistent with the main results (see Appendix Table A4 and Figure A2). I also replicated the main analyses using several transformations of the turnover measure. Results are similarly consistent with the main results (see Appendix Table A5 and Figure A3). I also estimate a version of this model that includes

state-by-year trends, while still allowing impacts vary based on time since QRIS adoption. Results indicate similarly positive impacts of QRIS adoption on teacher turnover, although point estimates are smaller in magnitude and less precisely estimated (see Appendix Table A6). However, this attenuation may be due the fact that QRIS adoption likely leads to a change in the outcome trend rather than having an immediate impact; therefore, this approach may incorporate some of the treatment effect into the state-specific trends.

Differential Impacts of QRIS Adoption Based on Year of QRIS Adoption

The results presented above consider the overall impact of QRIS adoption based on all states that adopted a QRIS between 2003 and 2018. However, changes in Head Start policy at the federal level during this period also broadly impacted the quality of care in Head Start. Although the differences-in-differences approach accounts for secular trends in Head Start quality, the presence of specific federal Head Start policies may have moderated the impacts of QRIS adoption. For example, the existence of federal requirements for the qualifications of Head Start teachers implemented after the 2007 HSPPS reauthorization may have influenced the extent to which the additional state-level accountability pressures and quality improvement support and incentives provided by states' QRIS led to further quality improvement.

To examine whether impacts of QRIS adoption varied based on the year of QRIS adoption, I conducted the primary difference-in-differences analysis separately for each year *t* between 2005 and 2016 after restricting the sample to states that adopted a QRIS year *t*, states that adopted a QRIS year ($t + 5$) or later, and states that never adopted a QRIS. I further restrict the sample to consider only the years 2003 through year ($t + 4$). This sample restriction provides a treatment-comparison contrast (QRIS vs. no QRIS) for states that adopted a QRIS up to four years after QRIS adoption. As shown in Figure 5, QRIS adoption had larger impacts on teacher

qualifications among states that adopted a QRIS relatively low early (i.e., before 2010). One explanation of this pattern of results is that QRIS adoption was sufficient to promote improvements in teacher qualifications before the federal Head Start requirements were implemented, but not after these requirements went into effect.

Variation in the Impacts of QRIS Adoption Based on the Use of Accelerated Rating Processes for Head Start

Next, I consider whether states' use of accelerated rating processes for Head Start programs moderates the effects of QRIS adoption on program quality and teacher turnover in Head Start. Results presented in Table 4 suggest that QRIS adoption had some impacts on teacher qualifications in states that used standard rating processes for Head Start programs, rather than using either reciprocity or alternative pathways to streamline Head Start programs' participation in the QRIS. The adoption of QRIS with standard rating processes for Head Start programs led to improvements in the percent of teachers with a bachelor's degree or higher (3.5 percentage points, $p < .05$). In contrast, results indicate that impacts of QRIS adoption on teacher turnover were driven by states with accelerating rating processes; impacts on teacher turnover among states with accelerated rating processes are larger by about 1.8 percentage points ($p <.10$) relative to states with standard rating processes.

In contrast, I find little evidence that impacts on process quality were larger in states that used standard rather than accelerated rating processes for Head Start programs. Results presented in the bottom panel of Table 4 suggest that impacts of QRIS adoption on instructional support were lower in state with reciprocity by about 0.45 SD ($p < .01$). However, similar differences were not observed based on states use of an alternative pathway. Differential impacts of QRIS adoption were also not observed for the emotional support or classroom organization domains.

I also conduct falsification tests analogous to the approach described above that estimate a model including indicators for placebo adoption of QRIS with each type of rating process. I find little evidence of pre-QRIS differences in outcome trends among state with standard rating processes, states with accelerated rating processes, and states without a QRIS (see Appendix Table B1). Nevertheless, states' decisions of how to design their QRIS are not randomly assigned and there may be other differences between states with standard and accelerated rating processes. Therefore, these results should be considered correlational rather than causal.

Evidence of Potential Mechanisms: Demand-side QRIS Factors

Next, I turn to the first set of exploratory analyses examining whether there is evidence to support specific hypothesized mechanisms of QRIS. First, I examine whether enrollment among higher-quality Head Start programs increased relative to lower-quality programs after QRIS adoption. Overall, as shown in Table 5, I find evidence that enrollment in higher-quality programs relative to lower-quality programs increased in the years following QRIS adoption. Specifically, programs with a relatively high proportion of teachers with a Bachelor's degree or above prior to QRIS adoption showed larger increases in enrollment after states' adoption of QRIS relative to programs with a lower proportion of teachers with a Bachelor's degree or above (a difference in log-enrollment of approximately 0.04; $p < 0.05$). Similarly, programs with relatively lower groups sizes prior to QRIS adoption also showed large increases in enrollment relative to programs with larger groups sizes. Results of falsification tests suggest that, when group size is used to define quality, these findings are not driven by differences in pre-QRIS trends among high-quality versus low-quality programs. Although there are some marginally significant differences in pre-QRIS trends among programs classified as high- versus low-quality

based on teacher qualifications, they are relatively small in magnitude relative to the estimated difference in the post-QRIS adoption period (see Appendix Table C1).

However, as shown in Table 6, I find little evidence that impacts on program quality in Head Start were larger the years in which families were more likely to be using quality ratings information to make child care decisions, based on information collected from *Google Trends*. In fact, results suggest a larger *negative* impact on teacher qualifications in the years when QRIS search popularity was higher. One potential explanation for this result is that increased demand for highly-qualified educators in other sectors of early childhood education – driven by the availability of QRIS ratings information – led highly-qualified teachers to switch from Head Start into other child care sectors. I return to this point below.

Evidence of Potential Mechanisms: Supply-side QRIS Factors

Next, I examine whether states' provision of supply-side supports and incentives is associated with larger impacts of QRIS adoption. Overall, I find little evidence to suggest that the provision of TA moderated the impacts of QRIS adoption on teacher qualifications or teacher turnover. Turning to the process quality outcomes, results provide some suggestive evidence that states' provision of more types of TA was positively associated with QRIS impacts on emotional support and, to a lesser extent, classroom organization CLASS scores. However, I do not observe similar impacts on instructional support (see Appendix Tables C2 and C3). Similarly, I find little evidence that states' use of financial incentives was associated with impacts of QRIS adoption on program quality in Head Start (see Appendix Tables C4).

Evidence Regarding ECE Compensation and Head Start Teacher Turnover

Finally, I examine whether effects of QRIS adoption on other sectors of early education and care systems – specifically, changes in compensation – might explain impact of QRIS

adoption on teacher turnover in Head Start[22]. In general, hourly wages for Head Start teachers exceeded those for other child care workers and preschool teachers (approximately $20 vs. $11)[23]. In most states, average child care/preschool wages increased the years following QRIS adoption among states with a QRIS; Head Start wages tended to show smaller increases or declines (see Appendix Figure D1). This led to a narrowing of the wage gap between educators in Head Start and educators in ECE programs in many states (see Appendix Figure D2). I conduct a preliminary difference-in-differences analysis examining the impact of QRIS adoption on relative hourly wages between child care/preschool and Head Start. Although estimates are imprecise, point estimates suggest QRIS increased wage parity (see Appendix Table D1).

As shown in Table 7, I observe that impacts on QRIS adoption on teacher turnover were larger when there was greater parity in the hourly wages of Head Start teachers and workers in other ECE programs – that is, when average child care and preschool wages were higher relative to Head Start wages. Point estimates suggest that impacts on overall teacher turnover were larger among states with higher child care/preschool wages relative to Head Start wages, although these impacts are imprecisely estimated. However, impacts of QRIS adoption on turnover due to teachers receiving higher compensation or benefits in other programs were significantly larger when child care/preschool wages were higher relative to Head Start wages. Specifically, a $1 increase in the average child care/preschool hourly wage (relative to the average Head Start hourly wage) was associated with 0.2 percentage point larger impact on turnover ($p < .10$).

[22] I exclude the District of Columbia from these analyses as I observe large fluctuations in the reported average hourly wage for Head Start teachers between 2004 and 2018.

[23] For Head Start, these figures are based on state-year level average hourly wages for Head Start classroom teachers (inflation adjusted to $2018), averaged across 2004 through 2018 and weighted by the number of Head Start teachers in each state-year. Data are from the Head Start PIRs. For child care/preschool, these figures are based on state-year levels average hourly wages for child care workers and preschool teachers (inflation adjusted to $2018), averaged across 2004 through 2018 and weighted by the number of Head Start teachers in each state-year. Data are from the Bureau of Labor Statistics Occupational Employment Survey (OES)

Results similarly indicate that states that experienced larger *reductions* in the wage gap following QRIS adoption also experienced greater increases in compensation-driven turnover. States that experienced a $1 larger increase in relative child care/preschool wages between the pre- and post-QRIS years also experienced larger impacts on turnover by 0.4 percentage points ($p < .05$). Although these findings are exploratory and correlational, results are consistent with the explanation that changes in the working conditions in non-Head Start sectors of ECE systems, possibly induced by QRIS adoption, affected the behaviors of Head Start teachers.

Discussion

The findings in the present study suggest that policymakers should use caution when considering QRIS as a tool to broadly improve the quality of care in Head Start programs. I find that the adoption of QRIS had little impact on either teacher qualifications or the quality of teacher-child interactions in Head Start. This is in in contrast to recent research that finds that states' adoption of QRIS led to overall improvements in educator qualifications among center-based programs (Herbst, 2018). However, I also find that impacts of QRIS adoption on teacher qualifications were larger among states that required Head Start programs to follow the standard set of QRIS rating procedures used by other center-based programs, as compared with to states that used accelerated or streamlined rating procedures for Head Start (e.g., automatic ratings). These findings are consistent with recent studies of QRIS that find evidence in some states, Head Start programs often enter in at the highest levels – due either to the use of accelerated rating processes or due to overlap between the federal Head Start and state QRIS standards – leaving little room for subsequent improvement (Herrmann et al., 2019; Kirby et al., 2019).

Despite lack of overall impacts on aspects of structural and process quality, findings suggest that Head Start programs were nevertheless impacted by states' QRIS adoption. I find

that QRIS adoption led to an increase in annual turnover rates among classroom teachers in Head Start of approximately 0.9 percentage points. There are several potential explanations for this increase in turnover. First, impacts on turnover could indicate that QRIS are working as intended to reduce the supply of low-quality care. Educators who do not meet the QRIS standards (e.g., in terms of professional qualifications or quality of instructional practice) may be exiting the market labor market. This phenomenon is similar, for example, to lower-rated teachers leaving the market in K-12 as a result of receiving low value-added scores (Dee & Wyckoff, 2015). However, the fact that the impact of QRIS adoption on Head Start teacher turnover is not accompanied by a corresponding positive impact on teacher qualifications or quality of teacher-child interactions provides suggests that lower-quality teachers may not be systematically leaving the Head start sector. Nevertheless, it is also possible that teacher qualifications and CLASS scores are not capturing the relevant quality measures; it may be that Head Start teachers are leaving programs because they don't meet other, unobserved quality dimensions.

Second, QRIS may increase job demands for teachers in Head Start programs. For example, QRIS ratings often increase teacher PD requirements, and require teachers to collect child assessment data (Tout et al., 2010). If these additional requirements are not accompanied by adequate supports (e.g., paid time to participate in PD), this may lead to worsened teaching conditions and increased turnover. However, a comparison of states with accelerated versus standard rating processes suggests that these pressures may not be the primary driver of increased turnover. The use of accelerated rating processes is intended to reduce the burden of Head Start programs' participation in the QRIS by streamlining the process used to achieve ratings (e.g., requiring programs to meet fewer standards; BUILD Initiative, 2018). However, I

find that impacts on teacher turnover are larger among states that use these accelerated rating processes relative to states that use standard rating processes.

A third explanation for the increase in teacher turnover in Head Start is that this impact is the result of QRIS-induced changes in other (non-Head Start) sectors of state ECE systems. Rates of QRIS participation among non-Head Start programs are generally high, particularly among center-based providers (BUILD Initiative, 2017). As non-Head Start programs typically have lower compensation levels and are subject to less intensive quality standards relative to Head Start programs (CSCCE, 2017), QRIS participation may have particularly large impacts in these programs. Indeed, Herbst (2018) finds that QRIS adoption led to an increase in certain aspects of quality as well as an increase in employee compensation in the center-based ECE sector. Furthermore, as Head Start programs represent only a fraction of the center-based sector, these impacts are likely to be driven by impacts on non-Head Start programs. It is possible that improvement in program quality (which could improve working conditions) or increases in compensation in these non-Head Start programs could lead to an increase in turnover as Head Start teachers switch to other programs within the center-based ECE sector. Exploratory analyses provide some support for this mechanism. I find that impacts of QRIS adoption on Head Start teacher turnover were larger when there was more parity between Head Start and child care and preschool wages. Additionally, I find that impacts were larger among states that experienced larger increases in salary parity between the pre- and post-QRIS adoption period. This suggests that impacts of QRIS adoption on quality or compensation in non-Head Start ECE programs may influence the behavior of teachers in the Head Start sector.

Limitations

I recognize several limitations in the present study. First and foremost, I do not have information about individual programs' participation in the QRIS. Due to this limitation, I am not able to examine whether impacts of states' adoption of QRIS are largest among the group of Head Start programs where were would expect to see the most impact: programs that are participating in the QRIS. Additionally, I am also able to observe the share overall share of Head Start programs (or non-Head Start programs) that participate in the QRIS in a given year, nor am I able to compare the features of programs that volunteer to participate in the QRIS relative to programs that do not participate. This limitation, in combination with the lack of information on program-level QRIS adoption, also does not allow me to more rigorously test for the possibility that Head Start programs that were not participating in the QRIS were nevertheless impacted by states' adoption of these systems, by examining impacts on programs who are not participating in the QRIS or in states with larger shares of participating programs.

An additional limitation is that I have information on Head Start programs, which may operate multiple centers, rather than individual Head Start centers. Recent research shows that there is considerable variation in programs quality across Head Start centers and classrooms (Sabol, Ross & Frost, 2019). Variation in quality across centers operated by the same program may mask variation in the impacts of QRIS adoption on centers operated by the same program. Finally, information on QRIS features is largely based on documents and resources collected after 2016; therefore, these analyses do not reflect changes in the features of QRIS over time (e.g., changes in the types of financial incentives or TA supports offered). Furthermore, I am not able to observe the frequency with which Head Start programs had access to supply-side resources and incentives. This may account for the lack of associations between supply-side features of QRIS (including TA and financial incentives) and impacts of QRIS adoption.

Conclusions

Quality Rating and Improvement Systems (QRIS) represent major state policy initiatives to unify and improve the quality of care in often-fragmented ECE systems. The present study provides insight into the impacts of these systems on Head Start programs, an important sector of ECE systems serving low-income children and families. Overall, results indicate that QRIS adoption did not broadly improved several aspects of structural and process quality in Head Start programs. These findings suggest the additional accountability pressures and incentives delivered by QRIS may not be sufficient to promote program quality in states that are already required to meet relatively stringent external monitoring requirements. Moreover, results indicate that QRIS adoption led to an increase in teacher turnover. This also raises the possibility of unintended consequences of QRIS adoption for these sectors of ECE systems. Future research should continue to examine the potential benefits – and costs – of incorporating programs that operate under federal policy contexts into states' QRIS.

References

Aikens, N., Bush, C., Gleason, P., Malone, L., & Tarullo, L. (2016). *Tracking quality in Head Start classrooms: FACES 2006 to FACES 2014*. OPRE Report 2016-95. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

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.

Barnett, W. S. (2003). *Better teachers, better preschools: Student achievement linked to teacher qualifications*. New Brunswick, NJ: National Institute for Early Education Research.

Barnett, W. S., Carolan, M. E., Fitzgerald, J., & Squires, J. H. (2012). *The state of preschool 2012: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research.

Barnett, W. S., & Friedman-Kruass, A. H. (2016). *State(s) of Head Start*. New Brunswick, NJ: National Institute for Early Education Research.

Barnett, W. S., Carolan, M. E., Squires, J. H., Clarke Brown, K., & Horowitz, M. (2015). *The state of preschool 2014. State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research.

Bassok, D., Dee, T. S., & Latham, S. (2019). The effects of accountability incentives in early childhood education. *Journal of Policy Analysis and Management*.

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.

Bassok, D., Magouirk, P., Markowitz, A. J., & Player, D. (2018). Are there differences in parents' preferences and search processes across preschool types? Evidence from Louisiana. *Early Childhood Research Quarterly, 44*, 43-54.

Bassok, D., Markowitz, A. J., Player, D., & Zagardo, M. (2017). *Do parents know "high quality" preschool when they see it*?. EdPolicyWorks Working Paper Series No. 54. University of Virginia.

Bloom, H. S., & Weiland, C. (2015). *Quantifying variation in Head Start effects on young children's cognitive and socio-emotional skills using data from the National Head Start Impact Study*. New York, NY: MDRC.

Boller, K., Paulsell, D., Del Grosso, P., Blair, R., Lundquist, E., Kassow, D. Z., ... & Raikes, A.

(2015). Impacts of a child care quality rating and improvement system on child care quality. *Early Childhood Research Quarterly, 30,* 306-315.

BUILD Initiative (2017). *Quality Compendium.* Available at: https://qualitycompendium.org.

BUILD Initiative (2018). *Toward integration: State approaches to Head Start partnering in Quality Rating Information Systems*. Available at: http://www.buildinitiative.org/Portals/0/Uploads/Documents/Resources/QRIS%203/Toward%20Integration_7-13-SC.pdf.

Cannon, J., Zellman, G. L., Karoly, L. A., & Schwartz, H. L. (2017). *Quality Rating and Improvement Systems for early care and education programs*. Santa Monica, CA: RAND Corporation.

Carneiro, P., & Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, *6*(4), 135-73.

Cascio, E. U., & Schanzenbach, D. W. (2013). The impacts of expanding access to high-quality preschool education. Brookings Papers on Economic Activity, 127.

Cassidy, D. J., Hestenes, L. L., Hansen, J. K., Hegde, A., Shim, J., & Hestenes, S. (2005). Revisiting the two faces of child care quality: Structure and process. *Early Education and Development*, *16*(4), 505-520.

Connors, M. C., & Morris, P. A. (2015). Comparing state policy approaches to early care and education quality: A multidimensional assessment of quality rating and improvement systems and child care licensing regulations. *Early Childhood Research Quarterly*, *30*(PB), 266–279.

Currie, J., & Neidell, M. (2007). Getting inside the "black box" of Head Start quality: What matters and what doesn't. *Economics of Education review, 26*(1), 83-99.

Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267-297.

Early, D., Barbarin, O., Bryant, D., Burchinal, M., Chang, F., Clifford, R., … Barnett, W. S. (2005, May). Prekindergarten in eleven states: NCEDL's multi-state study of pre-Kindergarten and Study of State-Wide Early Education Programs (SWEEP). University of North Carolina.

Herbst, C. M. (2018). The impact of quality rating and improvement systems on families' child care choices and the supply of child care labor. *Labour Economics*, *54*, 172-190.

Herrmann, M., Kirby, G., Deutsch, J., Wolfendale, C., Esposito, A. M., Caronongan, P. C., & Dragoset, L. (2019). *Quality ratings and system characteristics: Patterns in the round 1 Race to the Top-Early Learning Challenge states.* Evaluation Brief. NCEE 2019-4004. National Center for Education Evaluation and Regional Assistance.

Hestenes, L. L., Kintner-Duffy, V., Wang, Y. C., La Paro, K., Mims, S. U., Crosby, D., ... & Cassidy, D. J. (2015). Comparisons among quality measures in child care settings: Understanding the use of multiple measures in North Carolina's QRIS and their links to social-emotional development in preschool children. *Early Childhood Research Quarterly*, *30*, 199-214.

Hotz, V Joseph & Xiao, M. (2011). The impact of regulations on the suppy and quality of care in child care markets. *American Economic Review*, *101*(5), 1775–1805.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten

programs. *Early Childhood Research Quarterly*, *23*(1), 27-50.

Karoly, L. A., H. L. Schwartz, C. Messan Setodji, & A. C. Haas. (2016). Evaluation of Delaware Stars for Early Success: Final Report. Santa Monica, CA: RAND Corporation.

Kirby, G., Herrmann, M., Wolfendale, C., & Mraz-Esposito, A. Evaluation Brief. National Center for Education Evaluation and Regional Assistance.

La Paro, K. M., Thomason, A. C., Lower, J. K., Kintner-Duffy, V. L., & Cassidy, D. J. (2012). Examining the definition and measurement of quality in early childhood education: A review of studies using the ECERS-R from 2003 to 2010. *Early Childhood Research & Practice*, *14*(1), n1.

Ludwig, J., & Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly journal of economics*, *122*(1), 159-208.

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

Mitchell, A.(2009). Quality rating and improvement systems as the framework for early care and education system reform. Washington, DC: The Build Initiative

Mitchell, A. (2012). *Financial incentives in quality rating and improvement systems: Approaches and effects*. QRIS National Learning Network.

National Center on Early Childhood Quality Assurance (NCECQA) (2017a). *Financial incentives in the QRIS.* Available at: https://qrisguide.acf.hhs.gov/sites/default/files/QRIS_Financial_Incentives.pdf

National Center on Early Childhood Quality Assurance (NCECQA) (2017b). *QRIS Compendium 2016 Fact Sheet: Indicators of quality for ratings*. Available at: https://childcareta.acf.hhs.gov/sites/default/files/public/qris_indicators_quality_2016.pdf.

National Center on Early Childhood Quality Assurance (NCECQA) (2017c). *QRIS Compendium 2016 Fact Sheet: Staff qualifications, professional development, and supports*. Available at: https://childcareta.acf.hhs.gov/sites/default/files/public/qris_staff_qualifications_2016_0.pdf

Pianta, R., Downer, J., & Hamre, B. (2016). Quality in early education classrooms: Definitions, gaps, and systems. *The Future of Children*, *26*(2), 119-137.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system.* Baltimore, MD: Paul H. Brookes.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., ... & Ciarico, J. (2010). Head Start Impact Study. Final Report. Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services.

Roberts, J., N. L. Marshall, A. Tracy, S. Santaniello, M. G. Melia, H. Moore, S. Ellis, L. Kaufman, J. Tapper, S. Leibowitz, B. Comer, K. Gibbons, A. Glasgow, and K. Khlifi. (2016). Massachusetts Quality Rating and Improvement System (QRIS) Validation Study Final Report. Wellesley, MA: Wellesley Centers for Women.

Sabol, T. J., Hong, S. L. S., Pianta, R. C., & Burchinal, M. R. (2012). Can rating pre-k programs predict children's learning? *Science, 341*, 845–846.

Sabol, T. J., & Pianta, R. C. (2014). Do standard measures of preschool quality used in statewide policy predict school readiness? *Education Finance and Policy*, *9*(2), 116–164.

Sabol, T. J., & Pianta, R. C. (2015). Validating Virginia's quality rating and improvement system among state-funded pre-kindergarten programs. *Early Childhood Research Quarterly, 30,* 183–198.

Sabol, T. J., Ross, E. C., & Frost, A. (2019). Are all Head Start classrooms created equal? Variation in classroom quality within Head Start centers and implications for accountability systems. *American Educational Research Journal*.

Slot, P. L., Leseman, P. P. M., Verhagen, J., & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Childhood Research Quarterly*, *33*, 64–76.

Tout, K., J. Cleveland, W. Li, R. Starr, M. Soli, & E. Bultinck. (2016). The Parent Aware Evaluation: Initial Validation Report. Minneapolis, MN: Child Trends.

Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G., & Boller, K. (2010). *Compendium of quality rating systems and evaluations*. Washington, DC: U.S. Administration for Children and Families, Office of Planning, Research and Evaluation.

Tout, K., Zaslow, M., & Berry, D. (2006). Quality and qualifications: Links between professional development and quality in early care and education settings. *Critical issues in early childhood professional development*, 77-110.

US HHS Administration for Children & Families (2007*). Head Start Act, as amended December 12, 2007.*

US HHS Administration for Children & Families (2016a). *Head Start Program Performance Standards.*

US HHS Administration for Children & Families (2016b). *Designation Renewal System: DRS by the Numbers.*

QRIS Guide (2013). *Accreditation accepted for QRIS*. U.S. Department of Health and Human Services, Administration for Children and Families.

Walters, C. R. (2015). Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics*, *7*(4), 76-102.

Yazejian, N., & Iruka, I. U. (2015). Associations among tiered quality rating and improvement system supports and quality improvement. *Early Childhood Research Quarterly*, *30*, 255-265.

Tables and Figures

*Table 1***.** Head Start PIR and CLASS samples

	Head Start PIR sample		
	2003	2018	All years
Pct. lead teachers with Advanced degree	6.27	14.25	11.02
	(0.16)	(0.21)	(0.20)
Pct. lead teachers with Bachelor's degree	29.64	55.21	43.82
	(0.27)	(0.27)	(0.28)
Pct. lead teachers with Associate's degree	27.09	23.50	29.29
	(0.24)	(0.23)	(0.26)
Group size (# children per classroom)	18.16	16.71	17.05
	(3.70)	(3.03)	(3.27)
Teacher turnover (pct.)	14.71	19.35	16.06
	(0.18)	(0.19)	(0.18)
Total enrollment	501.03	451.80	506.02
	(661.84)	(547.69)	(659.87)
Total number of lead teachers	25.06	24.77	25.84
	(34.33)	(31.83)	(37.13)
Number of program-year observations	1,870	1,731	28,571
	Head Start CLASS sample		
Emotional Support	5.92	6.10	6.04
	(0.34)	(0.30)	(0.32)
Classroom Organization	5.48	5.82	3.74
	(0.48)	(0.39)	(0.42)
Instructional Support	3.03	2.99	2.94
	(0.71)	(0.58)	(0.58)
Number of grantee-year observations	334	394	2006

Note: For the Head Start PIR sample, table presents means of program-level characteristics with standard deviations in parentheses. For the Head Start CLASS sample, table presents means of grantee-level CLASS scores with standard deviations in parentheses.

Table 2. Falsification tests for impacts on structural quality, process quality, and teacher turnover

	Structural quality and teacher turnover			Process quality		
	Teachers with a B.A. or above (percent)	Group size	Teacher turnover (percent)	Emotional Support	Classroom Organization	Instructional Support
QRIS	-0.008	0.333	0.005	0.123	0.025	-0.149
	(0.023)	(0.203)	(0.009)	(0.109)	(0.109)	(0.103)
Pre-QRIS (1 year)	0.000	0.251+	-0.009	0.052	-0.070	-0.073
	(0.028)	(0.148)	(0.010)	(0.125)	(0.144)	(0.153)
Pre-QRIS (2 years)	-0.002	0.343*	-0.004	-0.021	-0.162+	0.041
	(0.022)	(0.160)	(0.008)	(0.124)	(0.088)	(0.205)
Pre-QRIS (3 years)	-0.004	0.211	0.001			
	(0.019)	(0.149)	(0.009)			
Pre-QRIS (4 years)	-0.008	0.265*	-0.005			
	(0.015)	(0.131)	(0.006)			
Observations	28,571	28,571	28,571	2006	2006	2006
p-value from *F*-test of QRIS leads	0.854	0.140	0.539	0.891	0.185	0.496

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Includes all fifty states and the District of Columbia. ** p<0.01, * p<0.05, + p<0.1.

Table 3. Impact of QRIS adoption on Head Start structural quality, process quality, and teacher turnover

	Structural quality and teacher turnover		Process quality		
	Teachers with a B.A. or above (percent)	Teacher turnover (percent)	CLASS: Emotional Support	CLASS: Classroom Organization	CLASS: Instructional Support
QRIS	-0.005	0.009*	0.106	0.081	-0.128
	(0.013)	(0.004)	(0.086)	(0.080)	(0.081)
Observations	28,571	28,571	2,006	2,006	2,006
States/ counties	54	54	51	51	51

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program/grantee covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Includes all fifty states and the District of Columbia.
** $p<0.01$, * $p<0.05$, + $p<0.1$.

*Table 4***.** Variation in the impact of QRIS adoption on Head Start structural quality, process quality, and teacher turnover based on the use of an accelerated rating process

Structural quality and teacher turnover:	Teachers with a B.A. or above (percent)		Teacher turnover (percent)			
QRIS	0.035*	0.035*	-0.000	-0.002		
	(0.016)	(0.015)	(0.006)	(0.006)		
QRIS*Accelerated process	-0.079**		0.018+			
	(0.029)		(0.010)			
QRIS*Reciprocity		-0.086*		0.022+		
		(0.036)		(0.013)		
QRIS*Alternative pathway		-0.066*		0.018+		
		(0.028)		(0.010)		
Observations	28,571	28,571	28,571	28,571		
State/counties	54	54	54	54		
Process quality:	CLASS: Emotional Support		CLASS: Classroom Organization		CLASS: Instructional Support	
QRIS	0.114	0.123	0.095	0.094	-0.105	-0.091
	(0.078)	(0.083)	(0.075)	(0.077)	(0.069)	(0.068)
QRIS*Accelerated process	-0.018		-0.033		-0.055	
	(0.095)		(0.079)		(0.093)	
QRIS*Reciprocity		-0.148		-0.110		-0.449**
		(0.160)		(0.194)		(0.166)
QRIS*Alternative pathway		-0.001		-0.001		0.036
		(0.105)		(0.089)		(0.090)
Observations	2006	2006	2006	2006	2006	2006
State/counties	51	51	51	51	51	51

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program/grantee covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Includes all fifty states and the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.

Table 5. Impact of QRIS adoption on relative enrollment in high vs. low-quality programs

	Outcome: ln(Total cumulative enrollment)		
QRIS	-0.027*	-0.021	-0.017+
	(0.012)	(0.012)	(0.009)
QRIS*High quality: Teachers with BA+ [a] *(in pre-QRIS years)*	0.037**		
	(0.013)		
QRIS*High quality: Group size [b] *(in pre-QRIS years)*		0.025+	
		(0.014)	
QRIS*High quality: Teacher turnover [c] *(in pre-QRIS years)*			0.017
			(0.021)
Observations	18,944	18,944	18,944

Note: Includes programs observed before and after QRIS adoption; excludes states without a QRIS. Standard errors in parentheses and clustered at the state level. All models include program and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Includes all fifty states and the District of Columbia.

** p<0.01, * p<0.05, + p<0.1.

[a] Programs were classified as high quality regarding teachers with a Bachelor's degree or above (BA+) as follows: First, in each year prior to QRIS adoption, I created an indicator for whether the share of teachers in a given program was above the median relative to other programs in the same state in the same year. Next, I classified programs as high quality if the program was above the median in the majority of pre-QRIS years in which the program was observed.

[b] Programs were classified as high quality regarding group size as follows: First, in each year prior to QRIS adoption, I created an indicator for whether the group size in a given program was below the median relative to other programs in the same state in the same year. Next, I classified programs as high quality if the program was below the median in the majority of pre-QRIS years in which the program was observed.

[c] Programs were classified as high quality regarding teacher turnover as follows: First, in each year prior to QRIS adoption, I created an indicator for whether the teacher turnover in a given program was below the median relative to other programs in the same state in the same year. Next, I classified programs as high quality if the program was below the median in the majority of pre-QRIS years in which the program was observed.

Table 6. Variation in the impacts of QRIS adoption on structural quality, process quality, and teacher turnover based on internet search popularity for the state QRIS

	Structural quality and teacher turnover			
	Teachers with a B.A. or above (percent)		Teacher turnover (percent)	
QRIS	-0.001		0.006	
	(0.016)		(0.005)	
QRIS*				
Search popularity	-0.146*	-0.194**	0.023	0.016
	(0.070)	(0.069)	(0.019)	(0.024)
Years since QRIS adopted				
0		-0.017		0.009
		(0.014)		(0.006)
1-2		0.001		0.003
		(0.016)		(0.005)
3-4		0.016		0.010
		(0.022)		(0.009)
5+		0.053		0.019*
		(0.034)		(0.008)
Observations	21,173	21,173	21,173	21,173
States/counties	40	40	40	40

	Process quality					
	CLASS: Emotional Support		CLASS: Classroom Organization		CLASS: Instructional Support	
QRIS	0.085		0.080		-0.086	
	(0.101)		(0.092)		(0.090)	
QRIS*						
Search popularity	-0.314	-0.394	-0.173	-0.212	-0.132	-0.286
	(0.255)	(0.312)	(0.230)	(0.278)	(0.275)	(0.327)
Years since QRIS adopted		0.126		0.090		-0.104
0		0.126		0.090		-0.104
		(0.117)		(0.118)		(0.112)
1		-0.065		-0.038		-0.026
		(0.124)		(0.132)		(0.125)
2		0.149		0.163		-0.155
		(0.115)		(0.106)		(0.096)
3+		0.141		0.103		-0.004
		(0.115)		(0.123)		(0.135)
Observations		1,675	1,675	1,675	1,675	1,675
States/counties		39	39	39	39	39

Note: Standard errors in parentheses and clustered at the state level. Search Popularity ranges from 0 to 100 within each state; 100 indicates peak search popularity. Search popularity is mean-centered for analysis. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program/grantee covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Includes all fifty states and the District of Columbia. ** p<0.01, * p<0.05, + p<0.1.

Table 7. Impact variation of QRIS adoption on teacher turnover based on salary parity

Outcome: Teacher turnover (percent)		
QRIS	0.022+	0.005
	(0.013)	(0.004)
QRIS*Hourly wage gap (CP - HS)	0.002	
	(0.002)	
QRIS*Change in hourly wage gap after QRIS adoption (CP - HS)[a]		0.003
		(0.002)
Observations	26,417	24,587
States/counties	53	47
Outcome: Teacher turnover due to compensation (percent)		
QRIS	0.015*	0.003
	(0.007)	(0.003)
QRIS*Hourly wage gap (CP - HS)	0.002+	
	(0.001)	
QRIS*Change in hourly wage gap after QRIS adoption (PS - HS) [a]		0.004*
		(0.002)
Observations	26,417	24,587
States/counties	53	47

Note: CP = child care/preschool weighted average; HS = Head Start. All dollar amounts adjusted to 2018 dollars. CP hourly wage calculated by taking average of CC and PS hourly wages, weighted by the number of individuals in each occupation. Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Analyses exclude the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.

[a] Models exclude states that did not adopt a QRIS between 2005 and 2018. The change in hourly wage gap calculated by the following: First, the difference between average child care/preschool and Head Start wages was calculated for each state/county and year. Second, the average wage gap for each state/county was calculated for the years before QRIS adoption (average pre-QRIS wage gap), and the years after QRIS adoption (average post-QRIS wage gap). Third, the change in the hourly wage gap was calculated by the difference between the average pre-QRIS wage gap and the average post-QRIS wage gap.

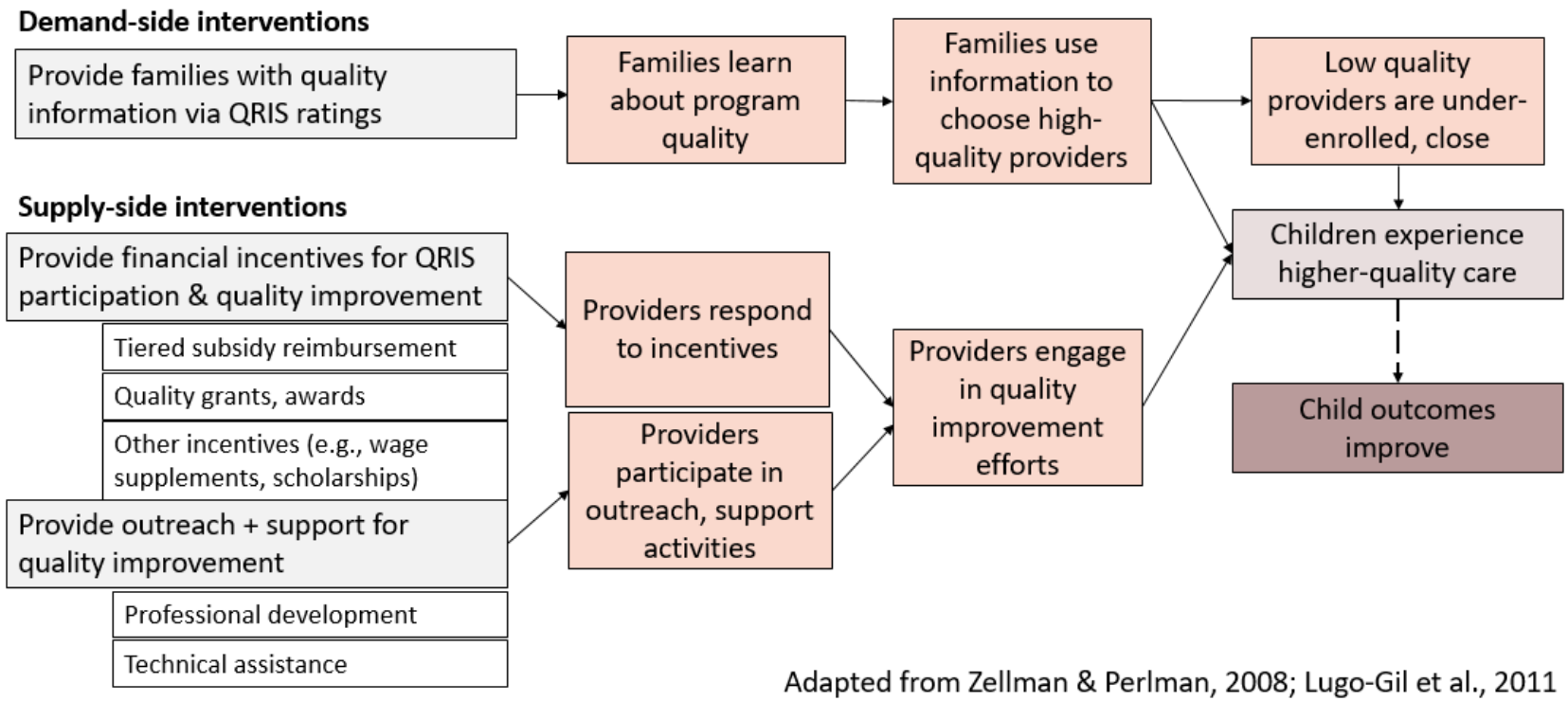


Figure 1. QRIS theory of action

Note: Figure shows the key components of most, but not necessarily all, state QRIS.

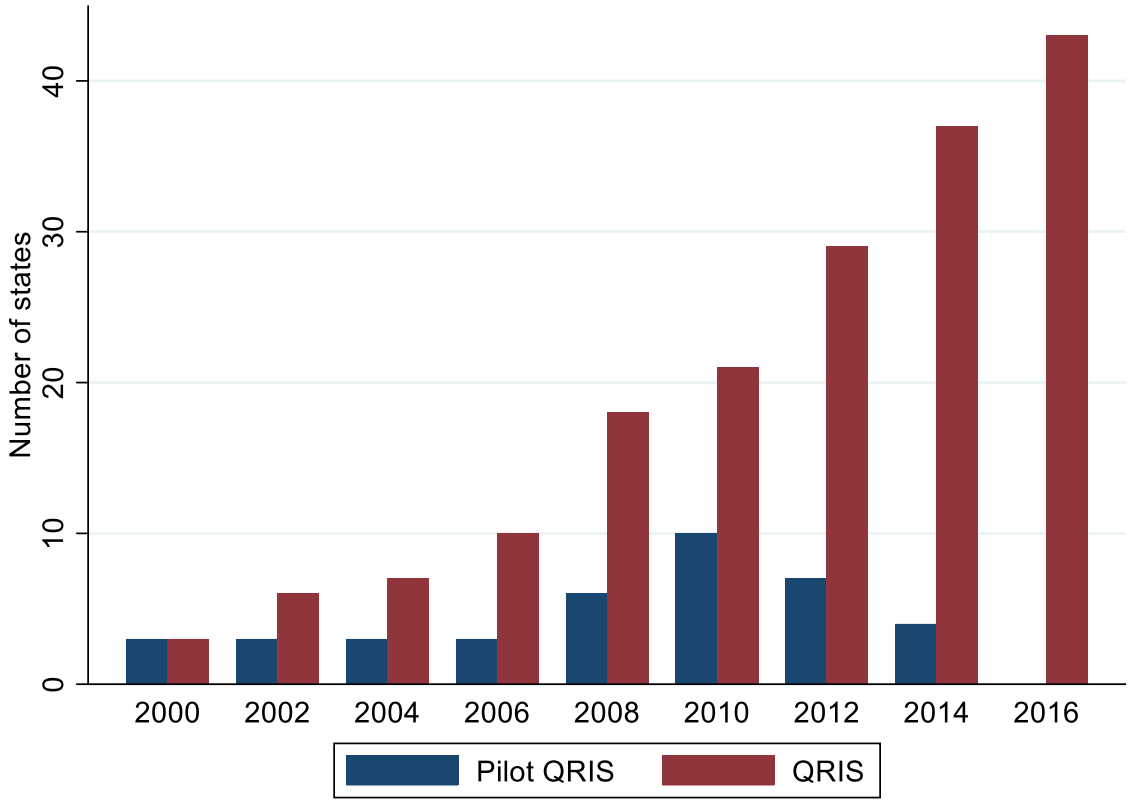
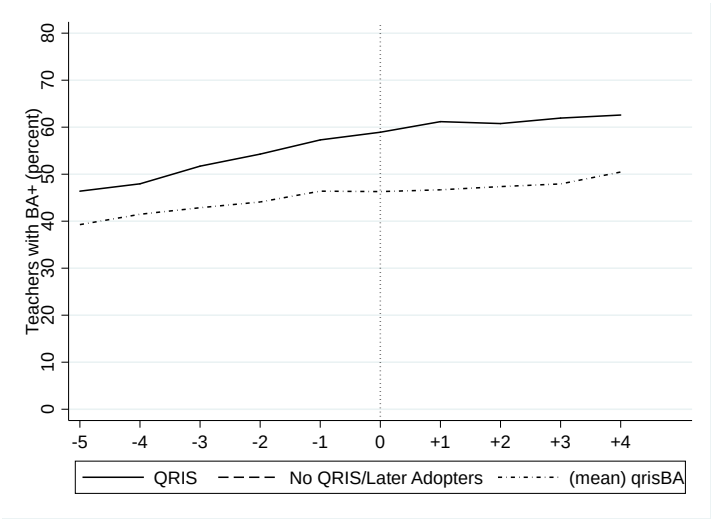
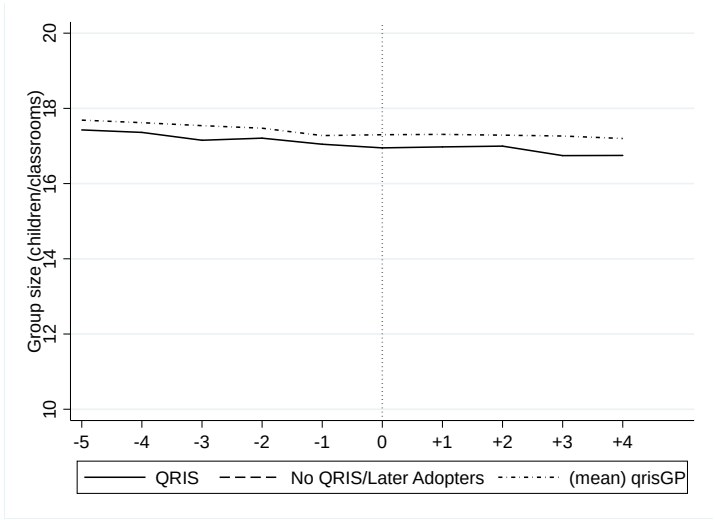


Figure 2. Adoption of QRIS from 2000-2016

Note: Figure shows the number of states that adopted a pilot QRIS/full QRIS in a given year. There are three counties that operate a QRIS in Florida. Figure shows when the first FL county adopted a pilot QRIS/full QRIS.



a) Teacher qualifications



b) Group sizes

c) Teacher turnover

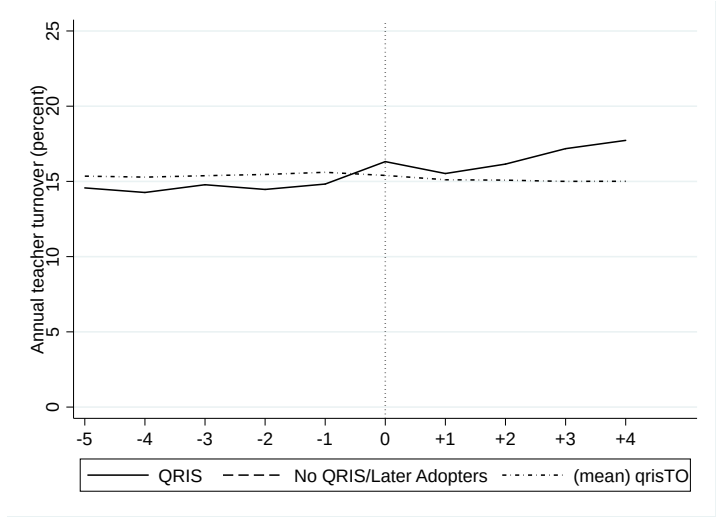
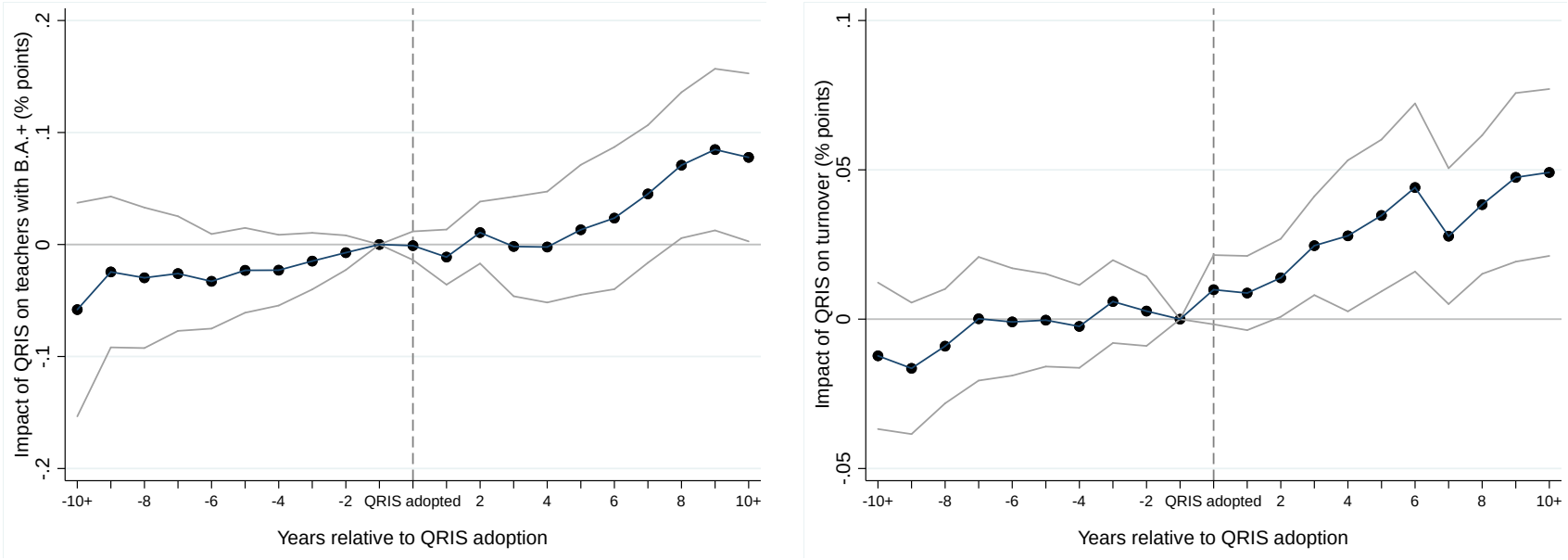


Figure 3. Structural quality and teacher turnover outcomes before and after QRIS adoption

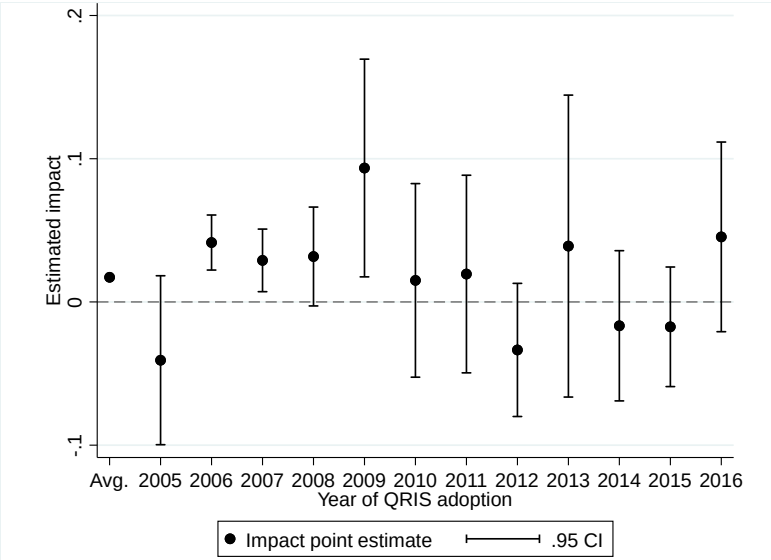


a) Impact on teacher qualifications

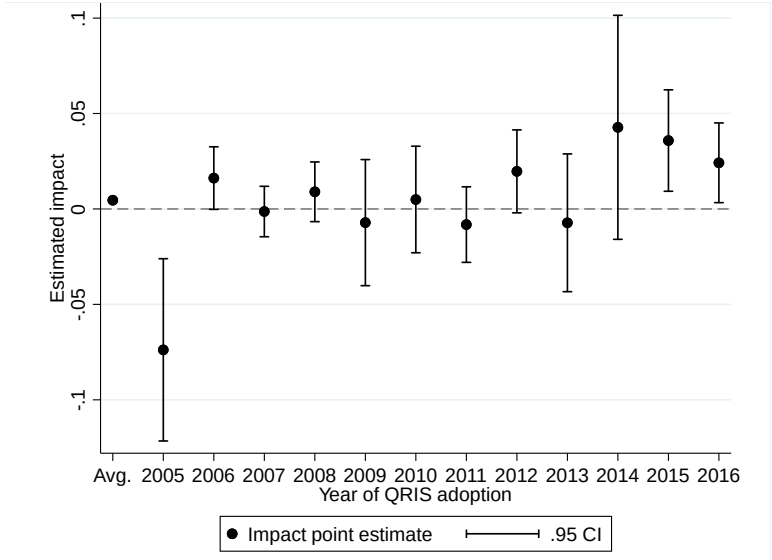
b) Impact on teacher turnover

*Figure 4***.** Differential impact of QRIS adoption based on time since system adoption

Note: Figures present impact estimates based on time since QRIS adoption with .95 confidence intervals.



a) Impact on teacher qualifications



b) Impact on teacher turnover

Figure 5. Differential impact of QRIS adoption based on year of QIS adoption.

Note: Large standard errors in 2005 are likely because only one small state (Vermont) and one Florida county adopted a QRIS in 2005. Figures present impact estimates based on time since QRIS adoption with .95 confidence intervals.

Appendix A: Supplementary Analyses, Overall Effect of QRIS Adoption

Table A1. Head Start CLASS Report Sample, 2012-2016

	2012		2018		All years
	ACF report	Analytic Sample	ACF report	Analytic Sample	Analytic Sample
Emotional Support	5.90	5.92	6.08	6.10	6.04
	(0.35)	(0.34)	(0.31)	(0.30)	(0.32)
Classroom Organization	5.45	5.48	5.80	5.82	3.74
	(0.49)	(0.48)	(0.39)	(0.39)	(0.42)
Instructional Support	2.98	3.03	2.96	2.99	2.94
	(0.70)	(0.71)	(0.55)	(0.58)	(0.58)
Number of grantee-year observations	388	334	462	394	2006

Note: Table presents means of grantee-level CLASS scores with standard deviations in parentheses.

Table A2. Evidence of changes in state prekindergarten enrollment, state prekindergarten spending, and federal Head Start funding to the state concurrent with QRIS adoption

	Percent of 4-year-olds enrolled in state-funded PreK	State spending per child in state-funded PreK	ln(Federal Head Start funding to the state)
QRIS	1.510	-44.776	-0.009
	(1.825)	(239.324)	(0.022)
Observations	860	858	864

Note: Observations at the state/county – year level. Standard errors in parentheses and clustered at the state level. All models include state/county fixed effects, year fixed effects, and control for state unemployment rate, female employment rate, and per-capita income.
** p<0.01, * p<0.05, $^{+}$ p<0.10

Table A3. Impact of QRIS adoption on Head Start structural quality, process quality, and teacher turnover, including group size

	Structural quality and teacher turnover			Process quality		
	Teachers with a B.A. or above (percent)	Group size	Teacher turnover (percent)	CLASS: Emotional Support	CLASS: Classroom Organization	CLASS: Instructional Support
QRIS	-0.005	0.098	0.009*	0.106	0.081	-0.128
	(0.013)	(0.161)	(0.004)	(0.086)	(0.080)	(0.081)
Observations	28,571	28,571	28,571	2,006	2,006	2,006
States/counties	54	54	54	51	51	51

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program/grantee covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Includes all fifty states and the District of Columbia.
** $p<0.01$, * $p<0.05$, + $p<0.1$.

Table A4. Impact of QRIS adoption on Head Start structural quality and teacher turnover, restricting the sample to states with a QRIS and without limited Head Start participation

	(1) Teachers with a B.A. or above (percent)	(2) Group size	(3) Teacher turnover (percent)
Excluding states that never adopted a QRIS			
QRIS	-0.012	-0.092	0.005
	(0.014)	(0.128)	(0.004)
Observations	25,921	25,921	25,921
States/counties	46	46	46
Excluding states with limited Head Start QRIS participation: AK, ND, TX, UT			
QRIS	-0.001	0.106	0.010*
	(0.014)	(0.170)	(0.004)
Observations	26,485	26,485	26,485
States/counties	50	50	50
Excluding states that adopted a QRIS before 2003			
QRIS	-0.008	0.207	0.006
	(0.014)	(0.180)	(0.004)
Observations	25501	25501	25501
States/counties	48	48	48
Excluding states that participated in RTT-ELC			
QRIS	-0.003	0.549*	0.018**
	(0.013)	(0.213)	(0.006)
Observations	14587	14587	14587
States/counties	34	34	34

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type Includes all fifty states and the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.

Table A5. Impact of QRIS adoption on Head Start teacher turnover using different outcome variable transformations

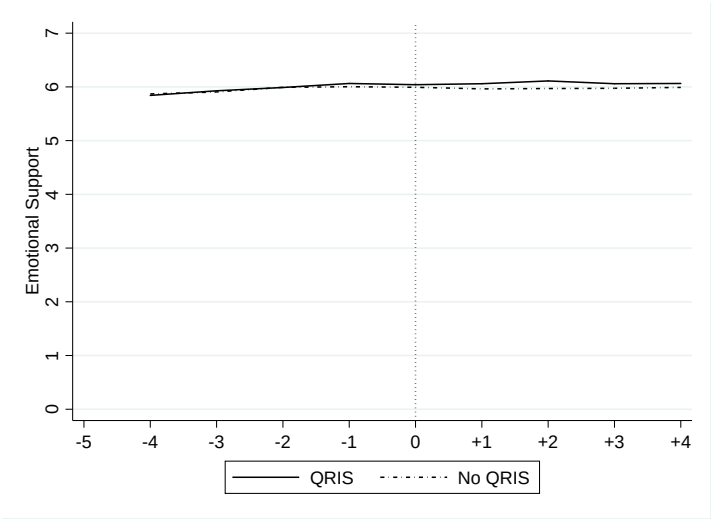
	(1)	(2)	(3)
	Outcome: Teacher turnover (percent)		
	Main specification	ln(Teacher turnover)	sqrt(Teacher turnover)
QRIS	0.009*	0.006+	0.007
	(0.004)	(0.003)	(0.005)
Observations	28,571	28,571	28,571
Num. states	54	54	54

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type Includes all fifty states and the District of Columbia. ** p<0.01, * p<0.05, + p<0.1.

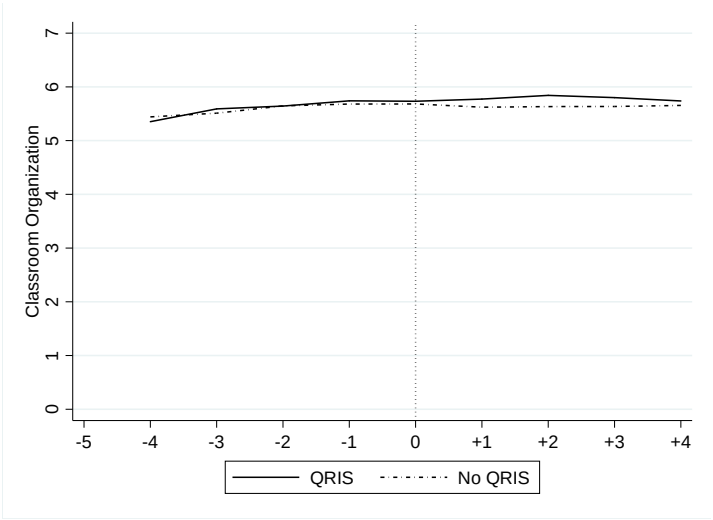
Table A6. Impact of QRIS adoption on Head Start structural quality and teacher turnover, including state-by-year trends

	(1) Teachers with a B.A. or above (percent)	(2) Group size	(3) Teacher turnover (percent)
QRIS:			
0-2 years post	-0.004	0.025	-0.001
	(0.007)	(0.175)	(0.004)
3-5 years post	-0.019	0.091	0.004
	(0.013)	(0.231)	(0.009)
6-8 years post	-0.022	0.326	0.007
	(0.021)	(0.295)	(0.013)
9+ years post	-0.029	0.453	0.015
	(0.026)	(0.361)	(0.014)
Observations	28,571	28,571	28,571
Num. states	54	54	54

Note: Standard errors in parentheses and clustered at the state level. All models include state/county, state/county-specific time trends, and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type Includes all fifty states and the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.



a) Emotional Support



b) Classroom Organization

c) Instructional Support

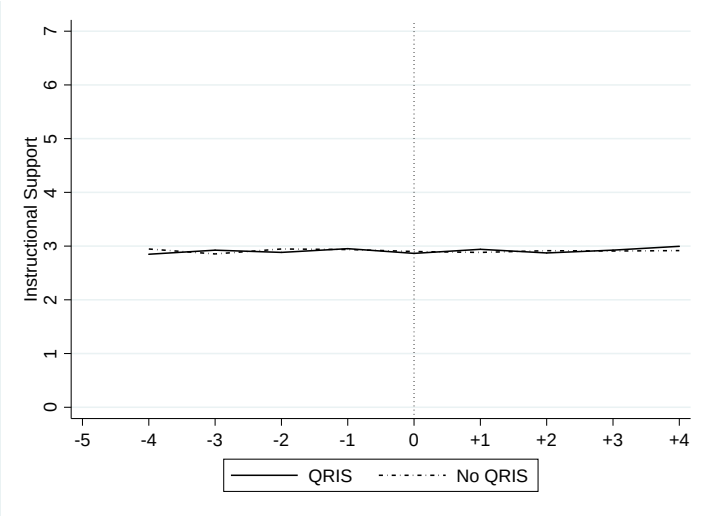
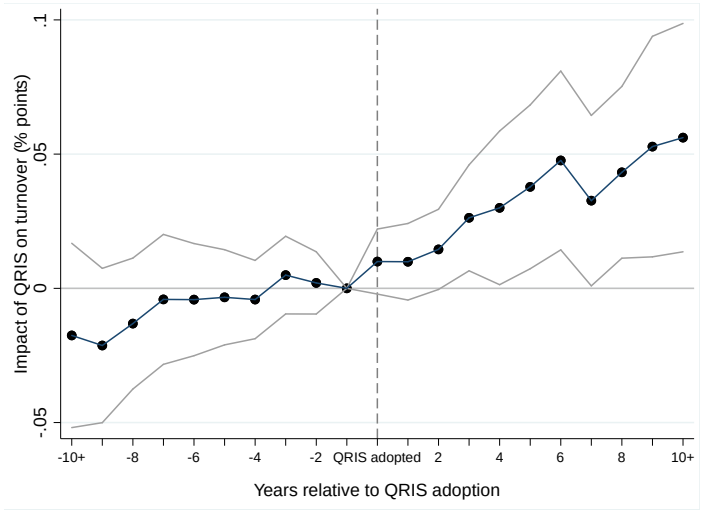
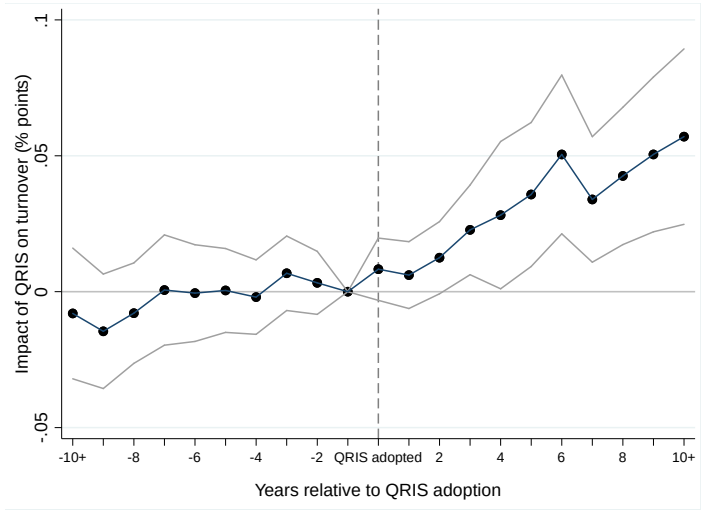


Figure A1. Process quality outcomes before and after QRIS adoption



a) Excluding states that never adopted a QRIS



b) Excluding states that adopted a QRIS pre-2003

c) Excluding states with limited QRIS adoption: AK, ND, TX, UT

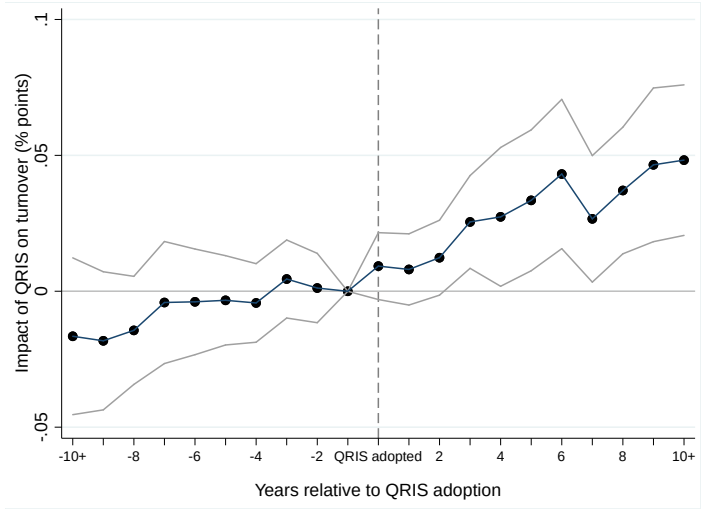
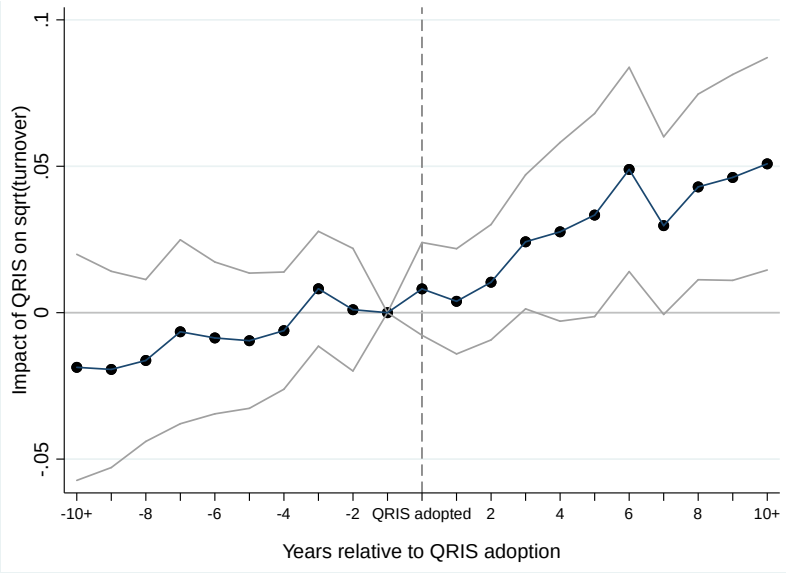
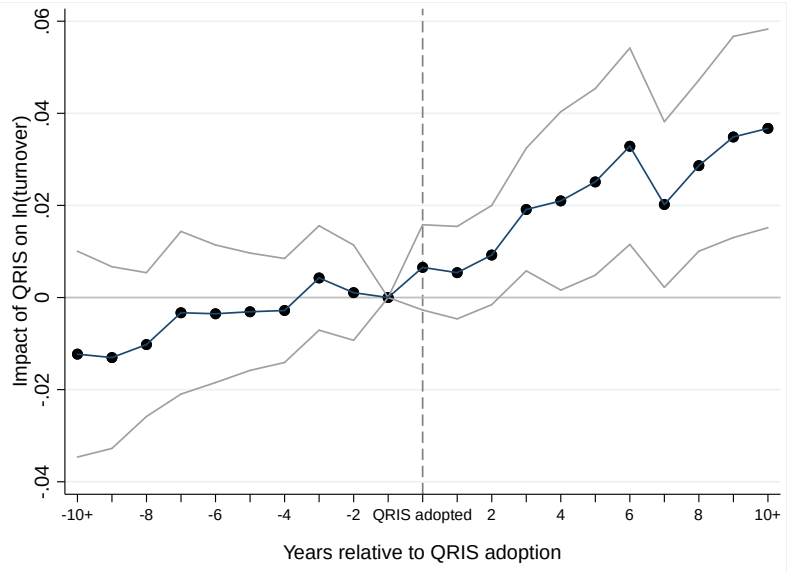


Figure A2. Impacts of QRIS adoption on teacher turnover based on years since QRIS adoption, using different sample specifications
Note: Figures present impact estimates based on time since QRIS adoption with .95 confidence intervals.



a) Outcome: sqrt(Teacher turnover)



b) Outcome: ln(Teacher turnover)

Figure A3. Impacts of QRIS adoption on teacher turnover based on years since QRIS adoption, using different outcome variable transformations

Note: Figures present impact estimates based on time since QRIS adoption with .95 confidence intervals.

Appendix B: Supplementary Analyses, Moderation of QRIS Impacts Based on Use of Accelerated Rating Processes

*Table B1***.** Falsification tests for impacts on structural quality and teacher turnover, including adoption of QRIS with and without accelerated rating process

	(1) Teachers with a B.A. or above (percent)	(2) Teacher turnover (percent)
QRIS without accelerated rating process	0.039	-0.005
	(0.032)	(0.010)
Pre-QRIS (1 year)	0.017	-0.011
	(0.038)	(0.012)
Pre-QRIS (2 years)	0.010	-0.005
	(0.031)	(0.009)
Pre-QRIS (3 years)	0.003	-0.005
	(0.025)	(0.010)
Pre-QRIS (4 years)	-0.005	-0.008
	(0.019)	(0.007)
QRIS with accelerated rating process	-0.047+	0.018+
	(0.025)	(0.010)
Pre-QRIS (1 year)	-0.015	-0.001
	(0.025)	(0.007)
Pre-QRIS (2 years)	-0.013	0.004
	(0.021)	(0.008)
Pre-QRIS (3 years)	-0.004	0.013
	(0.023)	(0.008)
Pre-QRIS (4 years)	0.005	0.004
	(0.021)	(0.008)

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates: race/ethnicity of children enrolled in programs operated by the grantee (percent black, percent white, percent Hispanic, percent American Indian/Alaska Native), age ranges of children in enrolled programs operated by the grantee (percent of 3 years old, percent of 4 years old), and home language of children enrolled in programs operated by the grantee (percent English). Includes all fifty states and the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.

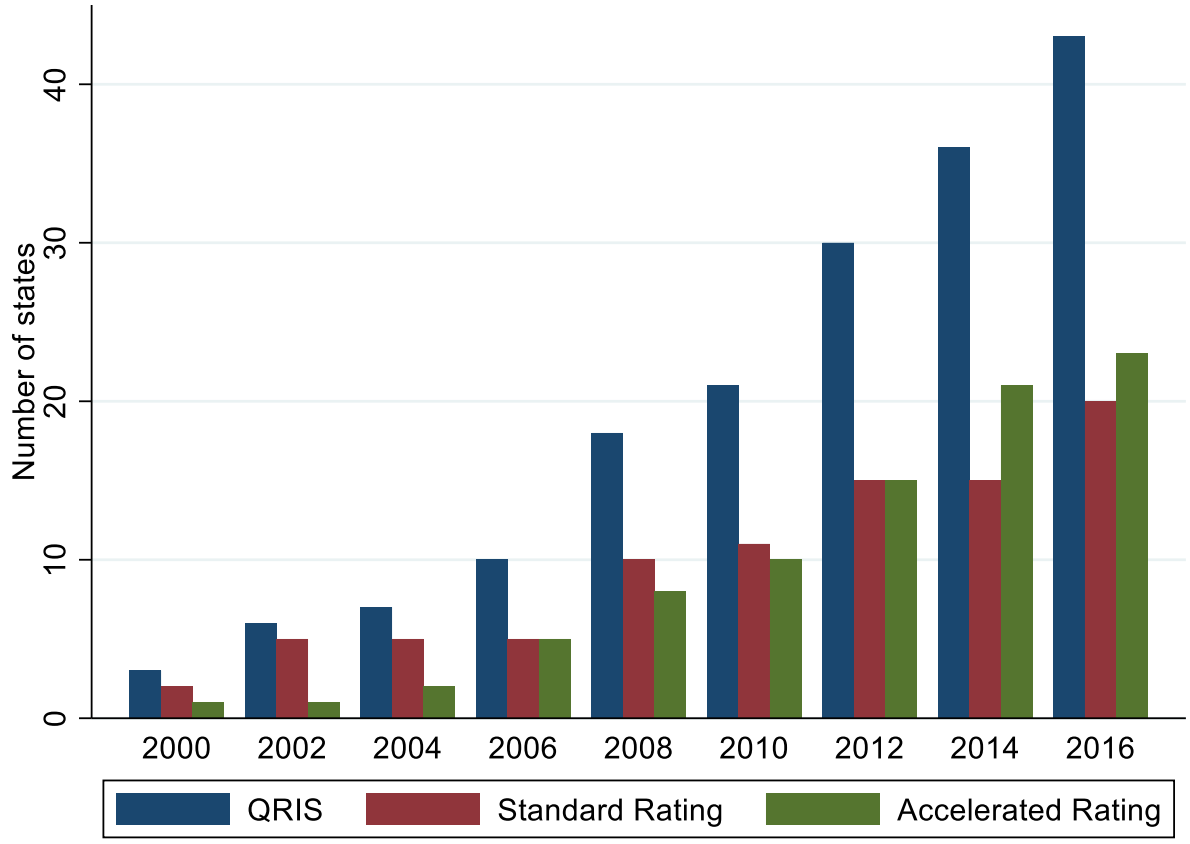


Figure B1. Adoption of QRIS with standard and accelerated rating processes from 2000-2016

Appendix C: Supplementary Analyses, Exploring Demand- and Supply-Side Mechanisms

Table C1. Falsification test for differential impacts of QRIS adoption on enrollment based on program quality

	(1)	(2)	(3)
	ln(Total cumulative enrollment)		
Quality measure:	Teachers with a B.A. or above (percent)	Group size	Teacher turnover (percent)
QRIS	-0.006	0.003	0.009
	(0.015)	(0.016)	(0.012)
Pre-QRIS (1 year)	0.018	0.023+	0.031*
	(0.013)	(0.013)	(0.012)
Pre-QRIS (2 years)	0.015	0.019	0.027*
	(0.010)	(0.012)	(0.012)
Pre-QRIS (3 years)	0.017	0.022	0.030*
	(0.012)	(0.013)	(0.013)
QRIS*High quality (*pre-QRIS)*	0.045**	0.031+	0.017
	(0.016)	(0.017)	(0.021)
Pre-QRIS (1 year)	0.023+	0.014	-0.003
	(0.012)	(0.016)	(0.010)
Pre-QRIS (2 years)	0.020+	0.015	-0.002
	(0.010)	(0.013)	(0.010)
Pre-QRIS (3 years)	0.014	0.005	-0.011
	(0.011)	(0.012)	(0.009)
p-value from *F*-test of QRIS leads	0.197	0.454	0.302

Note: Includes programs observed before and after QRIS adoption; excludes states without a QRIS. Standard error in parentheses clustered at the state level. All models include program and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type Includes all fifty states and the District of Columbia. ** p<0.01, * p<0.05, + p<0.1.

Table C2. Variation in the impact of QRIS adoption on structural quality, process quality, and teacher turnover based on number of types of QRIS technical assistance and financial assistance provided

	Teachers with a B.A. or above (percent)	Teacher turnover (percent)	CLASS: ES	CLASS: CO	CLASS: IS
Number of types of technical assistance provided					
QRIS	-0.064	0.021	-0.206+	-0.146	-0.283+
	(0.049)	(0.019)	(0.111)	(0.136)	(0.159)
QRIS*Number of TA types	0.018	-0.002	0.100**	0.070+	0.047
	(0.013)	(0.006)	(0.037)	(0.040)	(0.043)
N observations	28,121	28,121	1,973	1,973	1,973
N states/counties	53	53	50	50	50
Number of financial incentives					
QRIS	0.021	0.010+	0.232	0.148	0.000
	(0.014)	(0.006)	(0.178)	(0.158)	(0.137)
QRIS*Number of financial incentives	-0.018	0.000	-0.086	-0.057	-0.089
	(0.011)	(0.004)	(0.071)	(0.064)	(0.071)
N observations	25,800	25,800	1,843	1,843	1,843
N states/counties	52	52	49	49	49

Note: Standard error in parentheses clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type Includes all fifty states and the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.

Table C3. Variation in impacts of QRIS adoption on program quality based on state provision of technical assistance

	Teachers with a B.A. or above (percent)					Teacher turnover (percent)				
QRIS	-0.019	0.004	-0.021	-0.059^{+}	-0.005	0.020**	0.009	0.008	0.016	0.015*
	(0.028)	(0.055)	(0.021)	(0.035)	(0.019)	(0.007)	(0.016)	(0.011)	(0.011)	(0.007)
QRIS*Mentoring	0.030					-0.013				
	(0.040)					(0.012)				
QRIS*Coaching		-0.008					0.004			
		(0.057)					(0.017)			
QRIS*Consultation			0.022					0.007		
			(0.030)					(0.013)		
QRIS*PD advising				0.081^{+}					-0.003	
				(0.041)					(0.014)	
QRIS*Peer-to-peer					0.003					-0.005
					(0.037)					(0.012)
N observations	28,121	28,121	28,121	28,121	28,121	28,121	28,121	28,121	28,121	28,121
N states/counties	53	53	53	53	53	53	53	53	53	53

	CLASS: Emotional Support			CLASS: Classroom Organization			CLASS: Instructional Support		
QRIS	0.014	-0.062	-0.043	-0.030	-0.054	-0.007	-0.287*	-0.188^{+}	-0.218*
	(0.117)	(0.087)	(0.098)	(0.117)	(0.096)	(0.117)	(0.114)	(0.107)	(0.108)
QRIS*Mentoring	0.128			0.152			0.243^{+}		
	(0.161)			(0.142)			(0.138)		
QRIS*PD advising		0.302*			0.227^{+}			0.091	
		(0.137)			(0.133)			(0.146)	
QRIS*Peer-to-peer			0.263^{+}			0.135			0.151
			(0.142)			(0.140)			(0.143)
N observations	1,973	1,973	1,973	1,973	1,973	1,973	1,973	1,973	1,973
N states/counties	50	50	50	50	50	50	50	50	50

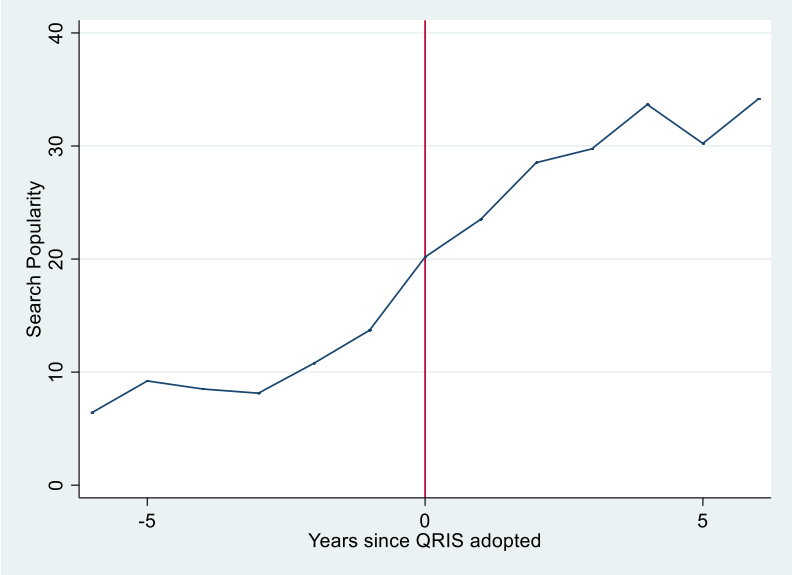
Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. Time-varying state covariates include the following: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Time-varying grantee covariates include the following: race/ethnicity of children enrolled in programs operated by the grantee (percent black, percent white, percent Hispanic, percent American Indian/Alaska Native), age ranges of children in enrolled programs operated by the grantee (percent of 3 years old, percent of 4 years old), and home language of children enrolled in programs operated by the grantee (percent English). ** p<0.01, * p<0.05, + p<0.1.

Table C4. Variation in impacts of QRIS adoption on program quality based on state provision of financial incentives

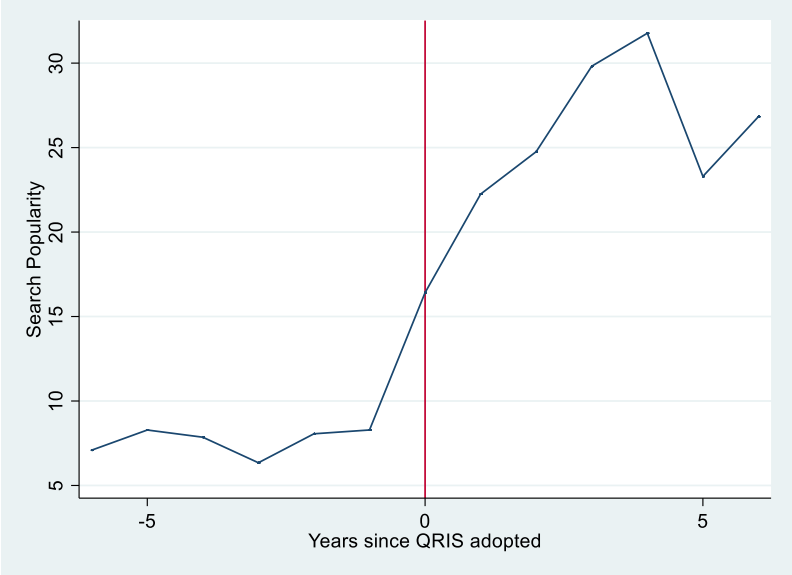
	Teachers with a B.A. or above (percent)			Teacher turnover (percent)		
QRIS	0.018	-0.018	0.018	0.005	0.018**	0.011*
	(0.013)	(0.018)	(0.013)	(0.005)	(0.006)	(0.005)
QRIS*Quality improvement grants	-0.053+			0.014		
	(0.027)			(0.010)		
QRIS*Quality awards/bonuses		0.026			-0.012	
		(0.027)			(0.009)	
QRIS*Scholarships			-0.052+			0.001
			(0.028)			(0.010)
N observations	25,800	25,800	25,800	25,800	25,800	25,800
N states/counties	52	52	52	52	52	52

	CLASS: Emotional Support		CLASS: Classroom Organization		CLASS: Instructional Support	
QRIS	0.155	0.090	0.079	0.082	-0.111	-0.108
	(0.144)	(0.151)	(0.134)	(0.137)	(0.122)	(0.127)
QRIS*Quality improvement grants	-0.165		-0.067		-0.092	
	(0.167)		(0.160)		(0.155)	
QRIS*Scholarships		-0.000		-0.066		-0.090
		(0.174)		(0.165)		(0.157)
QRIS*Scholarships						
N observations	1,843	1,843	1,843	1,843	1,843	1,843
N states/counties	49	49	49	49	49	49

Note: Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. Time-varying state covariates include the following: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Time-varying grantee covariates include the following: race/ethnicity of children enrolled in programs operated by the grantee (percent black, percent white, percent Hispanic, percent American Indian/Alaska Native), age ranges of children in enrolled programs operated by the grantee (percent of 3 years old, percent of 4 years old), and home language of children enrolled in programs operated by the grantee (percent English). ** p<0.01, * p<0.05, + p<0.1.



a) Search popularity, including all states



b) Search popularity, excluding states with a pilot QRIS

Figure C1. Popularity of internet searches for states' QRIS, based on years since QRIS adoption.

Appendix D: Supplementary Analyses, Exploring the ECE Compensation and Head Start Teacher Turnover

Table D1. Salary parity: Head Start, child care, and preschool

	Hourly wage						
	Head Start (HS)	Child care (CC)	Wage gap: CC vs. HS	Preschool (PS)	Wage gap: PS vs. HS	Child care/ Preschool (CP)	Wage gap: CP vs. HS
Summary statistics [a]							
Mean	$19.52	$11.28	-$8.24	$15.67	-$3.94	$12.91	-$6.60
N	750	750	750	750	750	750	750
Years				2004-2018			
Impact of QRIS adoption [b]							
QRIS	-$0.083	$0.181+	$0.264	-$0.309	-$0.234	$0.160	$0.236
	(0.238)	(0.108)	(0.250)	(0.255)	(0.368)	(0.136)	(0.303)
N	750	750	750	748	748	748	748
Years				2004-2018			

Note: CC = child care; HS = Head Start; PS = preschool; CP = child care/preschool. Observations at the state/county-year level, weighted by the number of Head Start teachers in each state/county and year. All dollar amounts adjusted to 2018 dollars. CP average hourly wage calculated by taking average of CC and PS hourly wages, weighted by the number of individuals in each occupation. All salary/hourly wage measures adjusted for inflation to 2018 dollars. State-level Head Start salary and wage information from Program Information Reports (PIRs). State-level preschool teacher and Child care wage information from the Bureau of Labor Statistics Occupational Employment Statistics (OES) Survey. Analyses exclude the District of Columbia. ** p<0.01, * p<0.05, + p<0.1.

[a] Results present means weighted by the number of Head Start teachers in each state/county in each year.

[b] Impacts of QRIS adoption based on results of estimating a difference-in-difference models including state/county fixed effects, year fixed effects, and time-varying state covariates; models estimated at the state/county-year level and include weights for the number of Head Start teachers in each state/county-in each year.

Table D2. Impact variation of QRIS adoption on teacher turnover based on salary parity

Outcome: Teacher turnover				
QRIS	0.021	0.007	0.018**	0.005
	(0.015)	(0.004)	(0.006)	(0.004)
QRIS*Hourly wage gap (CC – HS)	0.002			
	(0.002)			
QRIS*Change in hourly wage gap after QRIS adoption (CC - HS)		0.004		
		(0.003)		
QRIS*Hourly wage gap (PS - HS)			0.003+	
			(0.001)	
QRIS*Change in hourly wage gap after QRIS adoption (PS - HS)				0.003
				(0.002)
Observations	26,605	24,587	26,417	24,587
States/counties	53	47	53	47

Note: CC = child care; HS = Head Start; PS = preschool. All dollar amounts adjusted to 2018 dollars. Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Analyses exclude the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.

Table D3. Impact variation of QRIS adoption on teacher turnover due to compensation based on salary parity

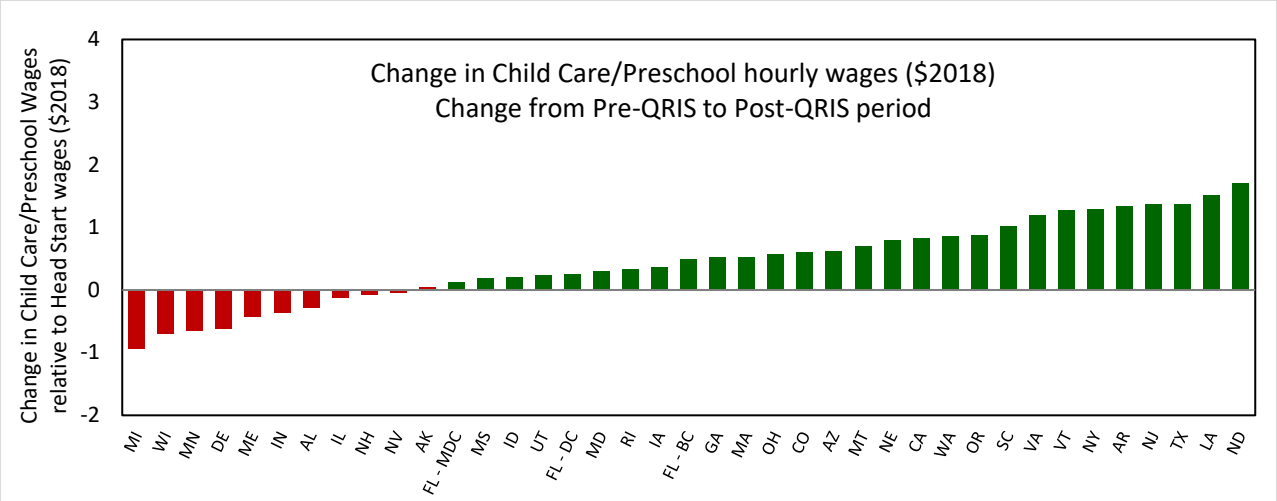
Outcome: Teacher turnover due to compensation				
QRIS	0.012	0.004	0.011**	0.002
	(0.007)	(0.002)	(0.003)	(0.003)
QRIS*Hourly wage gap (CC – HS)	0.001			
	(0.001)			
QRIS*Change in hourly wage gap after QRIS adoption (CC - HS)		0.003*		
		(0.002)		
QRIS*Hourly wage gap (PS - HS)			0.002**	
			(0.001)	
QRIS* Change in hourly wage gap after QRIS adoption (PS - HS)				0.002+
				(0.001)
Observations	26,605	24,587	26,417	24,587
States/counties	53	47	53	47

Note: CC = child care; HS = Head Start; PS = preschool. All dollar amounts adjusted to 2018 dollars. Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Analyses exclude the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.

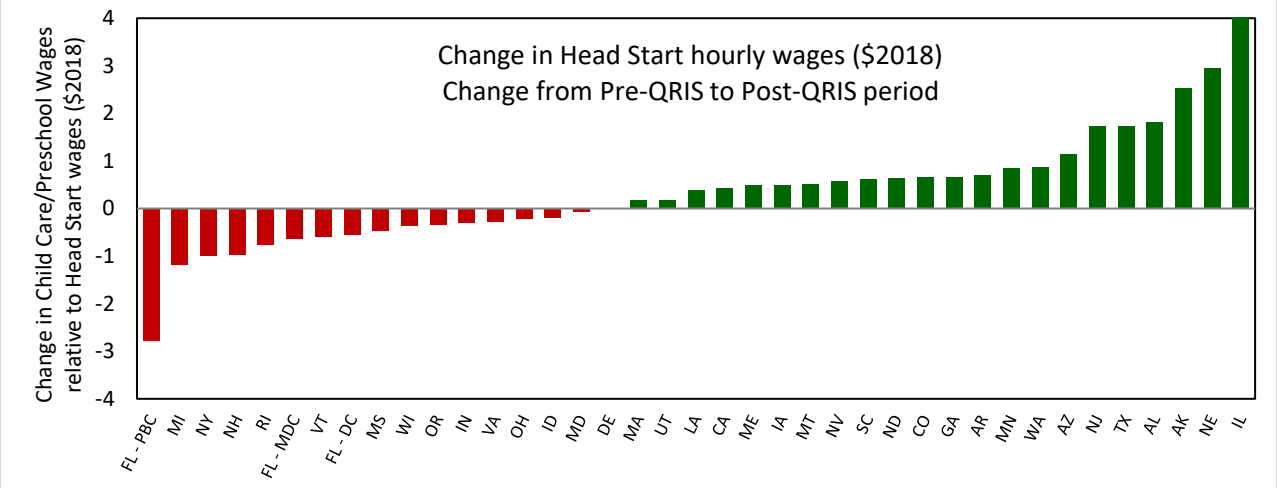
Table D4. Impact variation of QRIS adoption on teacher qualifications based on salary parity

Outcome: Percent of teachers with a B.A. or above		
QRIS	-0.014	-0.012
	(0.043)	(0.013)
QRIS*Hourly wage gap (CP – HS)	-0.002	
	(0.006)	
QRIS*Δ in Hourly wage gap (CP - HS)		-0.020*
Observations	26,417	24,587
States/counties	53	47

Note: HS = Head Start; CP = child care/preschool weighted average. All dollar amounts adjusted to 2018 dollars. CP hourly wage calculated by taking average of Child Care and Preschool hourly wages, weighted by the number of individuals in each occupation. Standard errors in parentheses and clustered at the state level. All models include state/county and year fixed effects and time-varying state and program covariates. State covariates include: annual state unemployment rate, annual female employment rate in the state, per capita income in the state, percent of four-year-olds enrolled in state-funded prekindergarten, state per-capita income, state spending per child enrolled in state-funded prekindergarten, and federal Head Start funding to the state. Program covariates include: race/ethnicity of enrolled children, ages of enrolled children, home language of enrolled children (percent English), and program type. Analyses exclude the District of Columbia. ** $p<0.01$, * $p<0.05$, + $p<0.1$.



a) Child Care/Preschool hourly wages



a) Head Start hourly wages

Figure D1. Change in hourly wages for Child Care/Preschool and Head Start.

Note: All dollar amounts inflation adjusted to $2018. The pre- vs. post-QRIS change in hourly wages calculated as the difference between the average hourly wage across the post-QRIS years and the average hourly wage across the pre-QRIS years. Green bars indicate that hourly wages *increased* after QRIS adoption. Red bars indicate that hourly wages *decreased* after QRIS adoption.

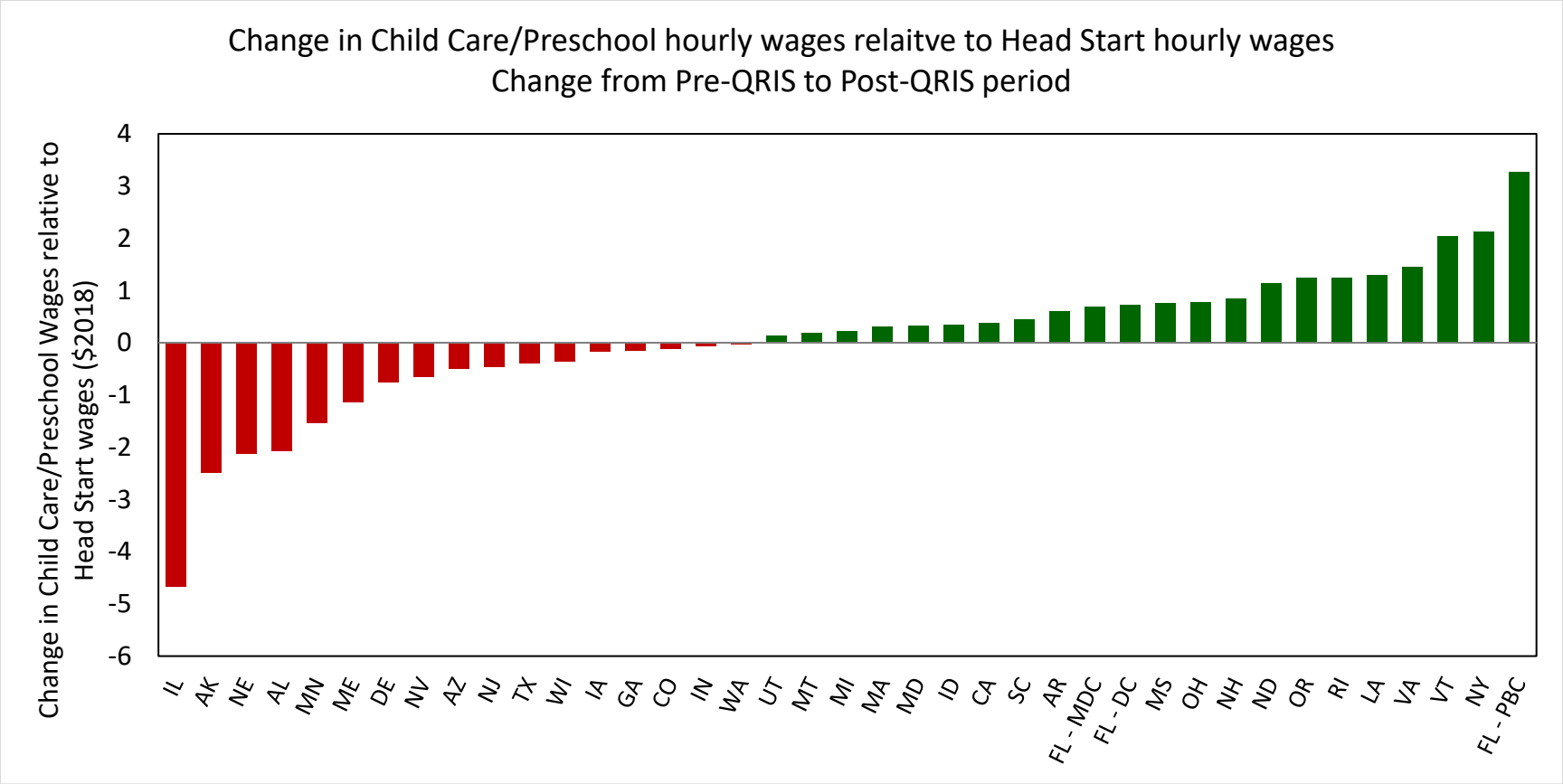


Figure D2. Change in hourly wage gap between Child Care/Preschool and Head Start.

Note: All dollar amounts inflation adjusted to $2018. The wage gap in each year calculated as the difference between Child Care/Preschool hourly wages and Head Start hourly wages. The pre- vs. post-QRIS change in wage gap calculated as the difference between the average wage gap among post-QRIS years and pre-QRIS years. Green bars indicate that the gap between Head Start and Child Care/Preschool wages *decreased* after QRIS adoption; i.e., that Child Care/Preschool wages *increased* relative to Head Start wages. Red bars indicate that the gap between Head Start and Child Care/Preschool wages *increased* after QRIS adoption; i.e., that Child Care/Preschool wages de*creased* relative to Head Start wages.