

Technical Report

MEASURES OF STUDENT NON-COGNITIVE SKILLS AND POLITICAL TOLERANCE AFTER TWO YEARS OF THE LOUISIANA SCHOLARSHIP PROGRAM

**EDUCATION
RESEARCH ALLIANCE**
.....
FOR NEW ORLEANS



SCHOOL CHOICE
DEMONSTRATION
PROJECT

Jonathan N. Mills, Albert Cheng, Collin E. Hitt,
Patrick J. Wolf, & Jay P. Greene

February 22, 2016

EducationResearchAllianceNOLA.com
UAedreform.org/school-choice-demonstration-project

**MEASURES OF STUDENT NON-COGNITIVE SKILLS AND POLITICAL
TOLERANCE AFTER TWO YEARS OF
THE LOUISIANA SCHOLARSHIP PROGRAM**

Jonathan N. Mills
Education Research Alliance for New Orleans,
Tulane University
New Orleans, LA
jmills10@tulane.edu

Albert Cheng
Department of Education Reform,
University of Arkansas

Collin E. Hitt
Department of Education Reform,
University of Arkansas

Patrick J. Wolf
Department of Education Reform,
University of Arkansas

Jay P. Greene
Department of Education Reform,
University of Arkansas

Louisiana Scholarship Program Evaluation Report #2
February 2016

School Choice Demonstration Project, University of Arkansas, Fayetteville, AR
Education Research Alliance, Tulane University, New Orleans, LA

Acknowledgments

We thank the Smith Richardson Foundation for financial support for this research. The content of the report is solely the responsibility of the authors and does not necessarily represent the views of the Smith Richardson Foundation, University of Arkansas, or Tulane University. We gratefully acknowledge Gema Zamarro, Anna Egalite, Lindsay Weixler, Martin West, and participants at the 40th Annual Conference of the Association for Education Finance and Policy for their extensive helpful comments on previous drafts and the Louisiana Department of Education for their cooperation and assistance with providing the necessary data to conduct the analyses.

Measures of Student Non-Cognitive Skills and Political Tolerance after Two Years of the Louisiana Scholarship Program

Abstract

This report examines the short-term effects of the Louisiana Scholarship Program (LSP) on students' non-cognitive skills and civic values. While a growing number of studies have evaluated K-12 school voucher programs along academic dimensions, few have focused on the development of non-cognitive skills and civic values. This study aims to address that gap by providing the first analysis of differences in self-reported measures of grit, locus of control, self-esteem, and political tolerance associated with the LSP. Using results from a phone survey of applicants to the program, we find little evidence of differences between LSP scholarship recipients and non-recipients. Nevertheless, diagnostics assessing the precision of our instruments to detect differences between subjects indicate that several of the scales measuring non-cognitive skills performed poorly in our sample. Moreover, our relatively low survey response rate of 11 percent raises concerns about the representativeness of our sample. Given these issues, we caution that our results are best understood as descriptive and not necessarily conclusive: they do not represent reliable estimates of the causal impact of the LSP on student non-cognitive skills and political tolerance.

Keywords: school vouchers, non-cognitive skills, conscientiousness, political tolerance, self-esteem, locus of control, civic values

Measures of Student Non-Cognitive Skills and Political Tolerance after Two Years of the Louisiana Scholarship Program

1. Introduction

While the majority of quantitative studies examining the effects of school choice programs on participating students have tended to focus on academic and attainment outcomes, a growing body of research suggests that other skills should also be evaluated (Mills, 2013). Specifically, studies have found that certain “non-cognitive” skills¹ such as conscientiousness and self-control are predictive of individual academic and workforce success (Almlund et al., 2011; Heckman & Kautz, 2012; Tangney, Baumeister, & Boone, 2004). More importantly, evidence suggests that different education environments can influence these skills (Cunha & Heckman, 2008; Dee & West, 2011; West et al., 2016). While two recent studies have examined the association between enrollment in charter schools and non-cognitive skills (Dobbie & Fryer, 2015; West et al., 2016), no such research exists for private school voucher programs. Moreover, there are only a handful of studies that have examined the impact of private school vouchers or scholarships specifically on student political tolerance (Fleming et al., 2014; Campbell, 2002; Howell & Peterson, 2002; Wolf et al., 2001). This is somewhat surprising, as school voucher programs give students the opportunity to enroll in private schools, many of which have missions explicitly emphasizing the importance of character-building and values formation in addition to academic achievement.

¹ As West et al. (2016) note, non-cognitive skill “has become a catch-all term for traits or skills not captured by assessments of cognitive ability and knowledge” (p. 1). In other academic disciplines, these skills are known as “character traits”, “personality traits”, and “soft skills”. Duckworth and Yeager (2015) note that this broad set of skills capture the extent to which individuals exercise sound judgment, goal-directed effort, and pursue healthy emotional relationships. Throughout this paper, we refer to this broad set of skills using the term “non-cognitive skills”.

This report aims to address the gap by providing the first descriptive analysis of differences in student self-reported measures of non-cognitive skills and political tolerance associated with the Louisiana Scholarship Program (LSP).

Specifically, we examine differences in student responses to several measures of non-cognitive skills and political tolerance collected through a phone survey conducted between November, 2014 and February, 2015. The non-cognitive measures included in our analysis are the Duckworth and Quinn (2009) Grit Scale, the Rotter (1966) Locus of Control Scale, the Rosenberg (1965) Self-esteem Scale, and the Sullivan, Pierson, and Marcus (1982) Political Tolerance Scale. We find no significant differences across all measures between students who received an LSP scholarship and those who did not.

There are, however, two caveats about our findings to underscore. First, internal reliability scores indicate the instruments we used to measure non-cognitive skills performed poorly in our phone survey sample and introduced a substantial amount of measurement error into our estimates. This fact increases the likelihood that we fail to detect differences between students who received an LSP voucher and those who did not, even if such differences actually exist. Thus, we caution that our results should not be interpreted as causal estimates of program effects.

Second, this work is best understood as a descriptive study focused on a subsample of LSP participants who chose to participate in our phone survey. While our sample, which represents 11 percent of those originally contacted, does not appear to differ strongly in terms of baseline characteristics from LSP applicants who chose not to respond, it is likely survey respondents differ from non-respondents in unmeasurable ways. Phone survey respondents, for example, were much more likely to have received an LSP scholarship compared to non-

respondents (72 percent of respondents received scholarships compared to 59 percent of non-respondents). The magnitude of this differential raises concerns that our sample of survey respondents may not be representative of the broader group of LSP applicants.

We proceed as follows. In the next section, we describe the existing literature examining the development of non-cognitive skills and civic attitudes in private school choice systems. We then outline the methodology used to estimate differences in these skills in the first group of students to experience the statewide expansion of the LSP. After that, we present our primary results and conclude with a summary of our findings along with implications for future research.

2. Non-Cognitive Skills, Civic Values, and School Choice

Evaluations of school voucher programs have generally focused on estimating voucher impacts on student achievement and attainment (Wolf, 2008; Wolf et al., 2013). The focus on academic outcomes is intuitive, as student achievement is linked to the economic success of individuals (Heckman, 2008) and countries (Hanushek & Woessman, 2009). Moreover, recent moves to test-based accountability systems in the United States have made measures of student achievement increasingly available to researchers (West et al., 2016).

At the same time, there is a growing body of research demonstrating that student non-cognitive skills—such as self-control and conscientiousness—are also predictive of short- and long-run life outcomes, even after controlling for academic achievement (Almlund et al., 2011; Heckman & Kautz, 2012). Furthermore, in the United States, there has been a long-standing view that schools play a crucial role in preparing children to participate in civic life (Dewey, 1916; Gutmann, 2003). These points suggest a need for school voucher evaluations to expand beyond measures of student achievement and attainment in assessing program effectiveness. In the following sections, we review literature supporting the importance of non-cognitive skills

and civic values, as well as existing evidence of the effectiveness of school choice programs in promoting these outcomes. The literature on the effects of voucher programs in developing these skills is quite limited, a finding which we argue supports the importance of our work.

The Importance of Non-Cognitive Skills and the Role of School Choice in Their Development

While evidence indicates positive associations between student academic achievement and later life outcomes (Heckman, 2008), recent empirical research has also demonstrated the independent role that non-cognitive skills play in short- and long-run life outcomes. For example, Tangney, Baumeister, and Boone (2004) find in an analysis of college students that measures of self-control are strongly related to college GPA, adjustment to college, and better emotional responses to stressful situations. More conscientious individuals are more successful in school, attain higher levels of education, are more likely to be employed, and earn higher incomes (Almlund et al. 2011; Borghans et al., 2008; Farkas, 2003; Lleras, 2008). Heckman and Kautz (2012) note that the positive associations between life outcomes and non-cognitive skills are to be expected, given that traits like conscientiousness and self-control tend to be particularly valuable in the workplace. Heckman, Stixrud, and Urzua (2006) also found evidence of negative associations between non-cognitive skills and the likelihood of incarceration and teenage pregnancy.

A growing number of studies have used measures of non-cognitive skills in educational settings, offering some evidence that educational environments can influence measures of these skills, if not the skills themselves. Dee and West (2011), using data from a nationally representative survey of children, find that smaller class sizes are positively related to measures of motivation and classroom engagement. Dobbie and Fryer (2015) find that students who win

admission via lottery to the Promise Academy middle school in Harlem, NY exhibit lower levels of grit relative to lottery losers, but not lower self-esteem. Similarly, West et al. (2014) find negative associations between attending over-subscribed charter schools and measures of student grit and conscientiousness. The authors of the latter two studies note, however, that it remains unclear if the negative results reflect actual changes in these skills or instead reflect the more demanding norms of their new schooling environments.

Unfortunately, the research estimating the effects of school voucher programs on the development of non-cognitive skills is limited. None of the existing experimental evaluations of school vouchers have included a non-cognitive component in their analysis beyond the proxies of high school attainment and college enrollment (Cowen et al., 2013; Chingos & Peterson, 2015; Warren, 2011; Wolf et. al., 2013;). This gap in the research base is surprising, given the specific emphasis placed by many private schools on discipline and non-academic outcomes. Private schools tend to have more rigorous and well-implemented discipline policies (Figlio & Ludwig, 2012). Catholic schools, in particular, are known for their emphasis on character building, in addition to academic achievement (Bryk, Lee, & Holland, 1993). Moreover, longitudinal studies find that private school attendance is associated with lower rates of teenage pregnancy, drug use, and engagement in related risky behaviors (Figlio & Ludwig, 2012; Mocan & Tekin, 2007). On the other hand, Elder and Jepsen (2014) find little evidence of positive effects of Catholic school attendance on student non-cognitive and behavioral outcomes, after accounting for selection bias.

In summary, the existing literature on non-cognitive skills development in schools of choice is limited. Among the 13 existing random assignment evaluations of school voucher programs in the United States, none have examined how vouchers impact the development of

students' non-cognitive skills. While existing evidence from charter school evaluations suggests the potential for negative impacts, there is strong reason to believe that these results are driven by different expectations across schools (Dobbie & Fryer, 2015; West et al., 2016).² Nevertheless, given increasing evidence of a link between non-cognitive skills and positive life outcomes, as well as the explicit focus of many private schools on such alternative educational outcomes, it is important, for the sake of completeness, to document the development of non-cognitive skills in school choice evaluations. The present study, which includes survey measures of non-cognitive skills, represents a first step in addressing this gap in the private school voucher literature.

Civic Skills and School Choice

Supporters of traditional public schools often argue that one of the primary benefits of democratically governed education is that it can educate children to promote civic values (Dewey, 1916; Gutmann, 2003; Henig, 1999). For example, Gutmann (2003) claims that “[a] central part of the historic mission of a democratically accredited school system is to educate citizens who are capable of sitting on juries, assessing public proposals (about schools, for example), exercising their rights, fulfilling their responsibilities, and seizing their opportunities to live a good life as they see fit” (p. 126). School choice, critics argue, will lead to a fractured system of education and will therefore fail to transmit collective values and foster social cohesion (Berliner & Biddle, 1996). Presumably, private schooling – especially faith-based

² This sentiment is shared by Tuttle and colleagues in their 2013 experimental evaluation of KIPP middle schools. They find that students randomly admitted to KIPP via oversubscription lotteries are more likely to report lying to their parents and losing their temper in school while also being significantly more likely to complete their homework on time. The authors note that these seemingly contradictory findings may be reflective of KIPP’s “no excuses” school environment, which places a strong emphasis on school discipline, hard work, and honesty.

private schooling – is at a comparative disadvantage relative to public schooling for fostering civic values such as political tolerance, which is a focus of this study.

It is unclear if private schools actually harm the development of political tolerance and other civic values. Although research has shown, for example, that religious dogmatism tends to be associated with lower levels of political tolerance (Gibson, 2010; Sullivan et al., 1982), faith-based private schools – and Catholic schools in particular – often actively emphasize the importance of the common good and respect for others, factors likely to foster tolerance (Candal & Glenn, 2012; Eisenstein, 2008; Scanlan, 2008).

Moreover, existing empirical evidence actually suggests that private schools tend to promote, or at least not harm, the development of politically tolerant individuals. Wolf (2005) examines the evidence on the effects of school choice and civic values in a systematic review, focusing on findings from experimental studies as well as rigorous quasi-experimental methods that approximate random assignment. In general, he finds 20 studies with 48 separate estimates of civic effects of private school choice meeting his selection criteria. Of the 48 total estimates, he finds only three indicating that private school choice negatively affects civic values. In contrast, 29 findings (60 percent) are either positive or contingently positive.³ Regarding political tolerance, Wolf cites seven studies that estimate the effects of private school attendance.⁴ Wolf finds private school attendance associated with higher levels of tolerance, although some studies found no difference between public and private school students.⁵ In conclusion, Wolf writes:

“The statistical record thus far suggests that private schooling and school choice rarely harms

³ Wolf (2005) categorizes a finding as contingently positive if it reports statistically significant positive findings for a type of private school rather than all private schools, and no negative impacts from any type of private school.

⁴ These studies were included because they either used random assignment (Campbell, 2002; Howell & Peterson, 2002; Wolf, Peterson, & West, 2001) or used quasi-experimental methods but were published in peer-reviewed journals.

⁵ Interestingly, one study found that attending a secular or Catholic private school was beneficial but attending a non-Catholic religious school undermined political tolerance (Campbell, 2001).

and often enhances the realization of the civic values that are central to a well-functioning democracy” (p. 237). New research on the Milwaukee Parental Choice Program, conducted after the Wolf review, reports that voucher participants demonstrated higher levels of political tolerance than matched public school students (Fleming et al., 2014).

The evidence presented in this section indicates a gap in the school voucher literature, whereby none of the existing voucher evaluations have examined non-cognitive skills development and few have considered political tolerance as an outcome variable. This paper is a first attempt at addressing this gap in the literature by describing differences in measures of non-cognitive skills and political tolerance among students who received and did not receive an LSP scholarship two years after the statewide expansion. The next section outlines the methodology used to study these topics.

3. Description of the Intervention

The Louisiana Scholarship Program is a statewide school voucher program available to moderate- to low-income students in low-performing public schools across the Pelican state. Student eligibility for the program is determined by family income—which must not exceed 250 percent of the federal poverty line—and the quality of the student’s previously attended public school. Income-eligible students must have attended a public school that was graded C, D, or F for the prior school year; be entering kindergarten⁶; or have been previously enrolled in the Recovery School District in order to be fully eligible for the program. In the program’s first year, 9,809 students were fully eligible applicants, with a majority of them located outside of Orleans parish.

⁶ Students applying for kindergarten were not required to have previously attended public schools with C, D, or F rankings.

The LSP was created by Act 2 of the 2012 Regular Session of the Louisiana Legislature and Senate. The voucher size is the lesser of the amount allocated by the state to the local school system in which the student resides and the tuition charged by the participating private school that the student attends. Average tuition at participating private schools ranges from \$2,966 to \$8,999, with a median cost of \$4,925, compared to an average total minimum foundation program per pupil amount of \$8,500 for Louisiana public schools.

The LSP was oversubscribed for the 2012-13 school year, with more applicants than scholarships available. To distribute the scholarships among the eligible applicants in the fairest and most efficient way possible, the program used a matching algorithm designed to take into account both the school preferences of families and the supply of available private schools. In particular, eligible applicants for the 2012-13 LSP cohort were allowed to submit up to five private school preferences. The LSP algorithm then attempted to match applicants with their most preferred school while giving former pilot program participants and new entrants from lower performing public schools slightly higher priority.⁷ Of the 9,809 eligible applicants for the 2012-13 cohort, 59 percent received LSP scholarships.

4. Methodology

This section introduces our methodology for investigating differences in non-cognitive skills and political tolerance between students who were awarded and not awarded an LSP voucher in the 2012-13 school year. We begin by describing the phone survey data collection process and then move on to a description of the non-cognitive and tolerance measures used in this study. The section concludes with a description of the final sample of survey respondents as well as a

⁷ A more detailed explanation of the LSP matching algorithm is provided in Appendix A.

comparison of our sample to the full population of eligible applicants in the 2012-13 program cohort.

Data Collection

Our study is based on the results of phone surveys of a subsample of the nearly 10,000 LSP-eligible applicants in the 2012-13 cohort, the first cohort participating in the program's statewide expansion. Our survey incorporates well-known scales designed to capture students' non-cognitive skills. Individual items were randomly ordered within scales to ensure that individual responses were not biased by question presentation order. In general, surveys lasted between 10 and 15 minutes. Our research team worked closely with an independent research group specializing in phone survey administration to complete data collection, which began on November 18, 2014 and concluded on February 7, 2015 after a total of 1,000 records were collected.⁸

Our research team provided the independent survey group with a randomly ordered list of LSP eligible applicants divided into two strata. The first strata consisted of students who received no exemptions in the LSP scholarship application process and therefore were more likely to have faced a scholarship lottery.⁹ The second strata included students who had participated in the New Orleans pilot program and students with special education exemptions.¹⁰

⁸ Upon contacting a household, surveyors first asked to speak with a parent or guardian to verify they had reached the intended family, described the purpose of the study, and requested consent to administer the survey to the child. After receiving consent, the surveyor asked to speak with the child, verified that the child's name matched the name on the intended record, and then administered the survey to the child. At the conclusion of the survey, the surveyor asked to speak again with the student's parent or guardian. The surveyor thanked the parent for their participation and provided the family with a toll-free number to call in case they had any additional questions about the study. Participants were offered no incentives, financial or otherwise, for participation in the study.

⁹ When seats were available, students with disabilities and multiple birth siblings (i.e., twins, triplets, etc.) were manually awarded scholarships to their desired school.

¹⁰ We excluded 159 students with severe disabilities from our call sample because their listed disabilities likely precluded their participation in the phone survey. Specifically, we excluded the following disability categories: Autism, Developmental Delay, Intellectual Disability (mild through severe), and Multiple Disabilities.

Our final survey sample consists of 999 students,¹¹ of whom 72 percent received an LSP scholarship.¹² This sample represents slightly more than 11 percent of the eligible applicants in the 2012-13 school year. The low response rate and sizeable difference in the percentage of students receiving scholarships between phone survey respondents and non-respondents present significant limitations for our analysis; a point we explore further in the sections below. Given these differences, we remind the reader that the analyses presented here are at best descriptive in nature and should not be interpreted as providing casual estimates of the program's impact.

These survey data have been merged with administrative data on student achievement and demographics provided by the Louisiana Department of Education (LDE). In addition, we have supplemented these data with information on school-level characteristics publicly available through the National Center for Education Statistics' Common Core of Data (CCD) and Private School Universe Survey (PSUS).

Measures of Non-Cognitive Skills and Civic Attitudes

This section describes four self-reported measures of non-cognitive skills and civic attitudes that are the basis for our study, chosen for their established links with later life outcomes. Each of the scales has been used in existing studies of school choice programs. For example, Dobbie and Fryer (2015) and West and colleagues (2016) include the Grit Scale in evaluations of the effects of charter school attendance on student outcomes. Dobbie and Fryer (2015) additionally include the Locus-of-control and Self-esteem Scales. Finally, the Political Tolerance Scale described in this section has been used in numerous studies of private schools (Wolf, 2005).

¹¹ Our final analytical sample excludes one of the original 1,000 respondents because the child's guardian later contacted the research team and asked that the child be removed from the study.

¹² As a comparison, 59% percent of applicants to the 2012-13 LSP cohort were awarded scholarships.

Initial diagnostics indicate that several of the scales perform poorly in distinguishing between students among our sample. For example, the internal reliability score, a measure capturing the ability of an instrument to consistently measure an unobserved latent trait, is particularly low for both the Grit and Locus-of-control Scales.¹³ The lower a scale's reliability score, the stronger the role random noise plays in the variation in scores we observe.¹⁴ Such measurement error is of particular concern to researchers as it tends to bias effect estimates towards zero, making it less likely that one could detect a program's effect even if it truly exists (Wooldridge, 2002). Moreover, scales based on self-reported surveys are less ideal measures of individual non-cognitive skills than behavioral assessments (Duckworth & Yeager, 2015; Egalite, Mills, & Greene, 2015). Given these limitations, we again recommend exercising caution when interpreting our results.

Grit. The first non-cognitive skill measured in this study is grit, or an individual's "perseverance and passion for long-term goals" (Duckworth et al., 2007, p. 1087). Our measure of grit is based on the 8-item Short Grit Scale developed by Duckworth and Quinn (2009), modified for young children.¹⁵ An individual's grit score is based on their average responses to eight five-point Likert scale items¹⁶ that include questions like "New ideas and projects sometimes distract me from previous ones" and "I am a hard worker".

¹³ These issues persisted even when we dropped seemingly problematic items from our survey and attempted to recalculate reliability scores. Rather than go with these subjectively adjusted scales, we choose instead to keep all items to maintain continuity with the original scales.

¹⁴ Some features of our design and implementation of the scales likely added noise to our final results. For example, none of the scales used in this study have been validated for phone surveys, nor have they been validated in populations as young as the study sample. The research team made minor changes to some of the survey items after consulting with the independent survey group to improve language clarity.

¹⁵ The adapted 8-item Grit Scale is available on Dr. Duckworth's website: <https://upenn.app.box.com/8itemgritchild>

¹⁶ Students are asked to choose among the following options: "Very much like you", "Mostly like you", "Somewhat like you", "Not much like you", and "Not like you at all".

Studies using different versions of the scale report that grit is predictive of several positive outcomes. Duckworth et al. (2007) report that grit is positively associated with career stability in a sample of adults, positively related to GPA among undergraduates at an elite Northeastern university, and is a better predictor of retention among West Point first-years than either a measure of self-control or an assessment administered by West Point. Duckworth and Quinn (2009) find that grit is positively related to student GPA, independent of IQ. On the other hand, two recent studies using the Grit Scale in evaluations of charter schools have identified negative relationships between charter school attendance and grit (Dobbie & Fryer, 2015; West et al., 2016). Both studies note, however, that the negative relationships may be driven in part by reference group bias resulting from differences in expectations across schooling environments. The Grit Scale has a 0.53 internal reliability score across our whole sample, with an internal reliability score of 0.52 for students in grades 2 through 6 and 0.58 among students in grades 7 through 12. The observed internal reliability scores for grit in our samples are substantially lower than the generally accepted threshold of 0.75 for internal reliability (Croker & Algina, 1986); however, this is not much lower than reliability scores observed in other school choice studies.¹⁷ Nevertheless, the low reliability scores suggests that much of the variation in scores we observe on this scale is due to measurement error,¹⁸ which further suggests estimates based on this scale will be biased towards null findings.

Locus of Control. The second scale included in our survey is the Locus of Control Scale developed by Rotter (1966), designed to capture the extent to which an individual believes rewards are the result of his or her own actions. We record an individual's locus of control based

¹⁷ For example, West et al. (2016) report a reliability coefficient of 0.64 for grit in their evaluation of Boston charter schools.

¹⁸ In general, a reliability score of .50 indicates that 50 percent of the variation in observed scores is due to noise.

on their responses to six four-point Likert scale items.¹⁹ The specific items are taken from the High School and Beyond Third Follow-up Survey administered by the U.S. Department of Education (1986) and include questions like “Good luck is more important than hard work for success” and “Every time I try to get ahead, something or somebody stops me”. The Locus of Control Scale has an internal reliability score of 0.47 across all phone survey respondents, with a score of 0.44 among students in grades 2-6 and 0.54 among students in grades 7 through 12. Reliability scores this low give us little confidence in our ability to detect the LSP’s role in producing meaningful variation in individual locus of control.

Self-esteem. We capture individual self-esteem levels using Rosenberg’s (1965) Self-esteem Scale. A respondent’s self-esteem score is calculated as their average response across 10 four-point Likert scale items. Each of the 10 items are designed to capture an individual’s view of their self-worth, including questions like “I am able to do things as well as most other people” and “I certainly feel useless at times”. In a 2003 review of studies using the Self-Esteem Scale, Baumeister et al. note that self-esteem is only moderately related to school performance, is a strong predictor of individual happiness, and is associated with a stronger likelihood of speaking up in a group, among other findings. The reported internal reliability score is 0.77 for the Self-Esteem Scale across all respondents, with a score of 0.73 reported for students in grades 2 through 6 and 0.83 reported for students in grades 7 through 12.

Political Tolerance. The final scale examined in this study attempts to capture participants’ civic attitudes by providing a measure of their political tolerance, defined as an individual’s willingness to permit the exercise of civil liberties by others with whom he or she disagrees. The political tolerance protocol developed by Sullivan et al. (1982) first asks

¹⁹ Individuals are asked to select among four responses to each question: “Strongly Disagree”, “Disagree”, “Agree”, and “Strongly Agree”.

individuals to identify a group that “has beliefs that [they] oppose the most” and then asks a series of questions regarding the level of political freedoms the individual would allow this group to enjoy. For example, individuals are asked if they “Strongly Disagree”, “Disagree”, “[are] Neutral”, “Agree”, or “Strongly Agree” that “The government should be able to secretly listen in on the telephone conversations” of their selected group. Unlike the three previous scales, the political tolerance scale was only administered to students without disabilities who were in grades 5 through 10 at baseline²⁰ due to the sensitive nature of the topic.²¹ The internal reliability score for this scale for this group of students is 0.77.

Sample Description

Data collection began in November of 2014 and continued for nearly four months until 999 records were collected. This group of respondents, representing 11 percent of all eligible LSP applicants in 2012, provide the basis for our primary analysis. This is a selective sample, based on the small proportion of families that opted into the phone survey. In particular, the low response rate raises some concerns that our results may not be representative of the broader population of eligible LSP applicants.

Table 1 presents descriptive statistics for several student characteristics collected at baseline for two groups of students: the students responding to our phone survey (columns 1 through 4) and all other eligible LSP applicants for the 2012-13 cohort (columns 5 through 8).

²⁰ These students should be in grades 7 through 12 as the time of survey administration unless they were held back during the time period examined.

²¹ In addition, the phone survey included a prompt before and during the questions noting, "If you are at all uncomfortable answering any of these questions, you may choose not to answer. That is completely ok."

The data presented in Table 1 are based either on student characteristics collected in the 2011-12 school year or from their LSP application.²²

In addition to describing the general demographics of our phone survey respondents, the comparisons presented in Table 1 offer insight into some of the issues facing our analysis. A first concern is the extent to which LSP recipients and non-recipients differ in our phone survey sample. Differences in baseline characteristics make it challenging to distinguish differences in outcomes associated with LSP participation from factors associated with these underlying differences in characteristics. Among phone survey respondents, LSP recipients are more likely to have participated in the New Orleans based LSP pilot program,²³ more likely to be female, offered slightly fewer school preferences, and are more likely to be enrolled in earlier grades than non-recipients. In addition, recipients responding to the phone survey tend to have performed worse on the state's assessments than non-recipients; however we only observe a statistically significant difference in math.²⁴ Given these differences, our preferred analytical model includes variables that control for these underlying characteristics.

²² Student grades, for example, were collected from a student's application. We would roughly expect these students to be two grades higher at the time of the survey if they were admitted to the grade applied for and progressed at a normal pace through grades. FRL and achievement data are only available for students in grades three through seven who took either the iLEAP or LEAP exams. Finally, a small percentage of students are missing information required to identify if they were living in a metropolitan statistical area at the time of application.

²³ Pilot program participants were given the highest priority status in the LSP matching algorithm (detailed in Appendix A).

²⁴ We can only observe student achievement for the subset of students who took the Louisiana assessments (iLEAP or LEAP) in grades three through seven at baseline. The size of this group is somewhat small, representing only a third of survey respondents; a factor contributing to the low statistical power of these analyses.

Table 1. *Characteristics of scholarship recipients and non-recipients, phone survey respondents vs. non-respondents*

	Phone Survey Sample				Other Eligible LSP Applicants (Non-Respondents to Phone Survey)				Difference in Differences
	N	Recipients:	Non-recipients:	Difference	N	Recipients:	Non-Recipients:	Difference	
		Mean	Mean			Mean	Mean		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
New Orleans Pilot Program	999	0.35	0.02	0.33***	8530	0.30	0.02	0.28***	0.05***
Female	999	0.55	0.46	0.09***	8530	0.51	0.49	0.02*	0.07**
African American	999	0.84	0.84	0.00	8530	0.87	0.87	0.00	0.00
White	999	0.08	0.11	-0.02	8530	0.08	0.08	0.00	-0.02
Hispanic	999	0.03	0.04	0.00	8530	0.02	0.03	-0.01*	0.00
FRL Eligible, Non-Pilot ^a	291	0.93	0.93	0.00	2920	0.93	0.94	-0.02*	0.02
Living in MSA ^b	921	0.95	0.96	-0.01	7588	0.96	0.97	-0.01*	0.00
# of School Preferences Listed	999	1.82	2.36	-0.54***	8530	1.92	2.25	-0.33***	-0.21**
Grade at Application									
Kindergarten	999	0.22	0.18	0.04	8530	0.24	0.18	0.06***	-0.02
First	999	0.15	0.10	0.05**	8530	0.14	0.10	0.05***	0.00
Second	999	0.12	0.08	0.04**	8530	0.13	0.07	0.06***	-0.02
Third	999	0.16	0.11	0.05**	8530	0.11	0.10	0.01*	0.04*
Fourth	999	0.08	0.11	-0.03	8530	0.09	0.11	-0.02***	-0.01
Fifth	999	0.09	0.08	0.01	8530	0.07	0.09	-0.01*	0.02
Sixth	999	0.08	0.10	-0.01	8530	0.08	0.09	-0.01**	0.00
Seventh	999	0.04	0.09	-0.04**	8530	0.06	0.10	-0.04***	0.00
Eighth	999	0.02	0.08	-0.06***	8530	0.03	0.08	-0.05***	-0.01
Ninth	999	0.03	0.06	-0.04**	8530	0.03	0.06	-0.03***	-0.01
Tenth	999	0.01	0.02	-0.01	8530	0.02	0.02	-0.01***	0.00
Achievement Scores ^c									
Math	355	-0.58	-0.42	-0.16*	3330	-0.54	-0.49	-0.05	-0.11
ELA	356	-0.50	-0.38	-0.12	3328	-0.46	-0.40	-0.06*	-0.06
Science	355	-0.61	-0.43	-0.18	3322	-0.60	-0.50	-0.10***	-0.08
Social Studies	355	-0.51	-0.37	-0.14	3322	-0.51	-0.44	-0.07*	-0.07

*** - $p < .01$, ** - $p < .05$, * - $p < .10$

Note. ELA is English Language Arts. [a] FRL eligibility for non-New Orleans LSP Pilot Program. Student FRL status is only available for students appearing in the state's testing data. [b] Data on Metropolitan Statistical Area (MSA) is taken from the American Community Survey. [c] Student achievement on the iLEAP or LEAP exams has been standardized within subject and grade to the state's testing distribution. Difference in means tests presented in columns 4, 8, and 9 are based on heteroscedasticity robust standard errors. Source. Authors' calculations.

Columns 5 through 8 compare the baseline characteristics of LSP recipients and non-recipients to those in the target population who did not participate in our phone survey. By comparing columns 2 and 3 with columns 6 and 7 we find that our survey respondents do not differ greatly from non-respondents on most baseline characteristics. In addition, we observe similar patterns of differences among non-respondents between scholarship recipients and non-recipients: recipients are more likely to have participated in the New Orleans pilot program, are more likely to be female, offered fewer school preferences, applied for earlier grades, and performed slightly worse on state assessments than non-recipients. Among those who did not respond to the phone survey, students who received a scholarship were slightly less likely to be Hispanic, eligible for free or reduced-priced lunch, or living in a metropolitan area. Column 9 examines how strongly the recipient/non-recipient characteristic differentials differ between survey respondents and non-respondents. While we observe some differences, for the most part, the difference-in-difference estimates are small.

5. Results and Discussion

In the following sections, we present the primary results from our analyses examining differences in measures of students' non-cognitive skills and political tolerance. The evidence presented here largely suggests that the two groups of students did not differ across any of the four measures of interest two years after initial LSP scholarship assignment. We wish to emphasize, however, that these are neither causal estimates nor do we place much confidence in them due to the measurement issues described earlier.

Descriptive Analysis

To better understand the measures of students' non-cognitive skills that were collected, we check for correlations between each of the non-cognitive measures and students' learning gains in math and ELA. Table 2 presents pairwise correlations between measures of the three non-cognitive skills and political tolerance along with four estimates of student achievement growth from the 2011-12 to the 2012-13 school year.²⁵ We expect to observe a positive relationship between these two sets of measures, which is largely confirmed by the data. Panel A presents results for the full set of respondents with complete responses for all measures, excluding political tolerance. Panel B presents results for a subset of students who additionally provided responses for the Political Tolerance Scale.²⁶ We include achievement score gains in Table 2 to examine the relationship between the included non-cognitive skills measures and student achievement gains; however in doing so, we have substantially restricted the sample for which we can estimate these relationships. Nevertheless, the relationships observed in Table 2 among the non-cognitive skills measures generally hold in the full sample of survey respondents.

²⁵ Keeping in line with the work of West and colleagues (2016), we calculate mean performance gain as the average residual resulting from a regression of standardized achievement in 2012-13 on a cubic function of achievement in 2011-12.

²⁶ Due to the sensitive nature of the items on the Political Tolerance survey, we only administered the scale to students in grades 7 through 12 in the fall of 2014 who did not indicate a disability on their original LSP application.

Table 2

Correlation matrices of non-cognitive skills, tolerance, and achievement growth measures, by age group

	Grit	Locus of Control	Self-Esteem	Political Tolerance	Residual Math Gain	Residual ELA Gain	Residual Science Gain
Panel A: Phone survey sample with complete responses (N=229)							
Locus of control	0.43***						
Self-Esteem	0.37***	0.53***					
Residual Math Gain	-0.02	-0.02	-0.03				
Residual ELA Gain	0.06	0.08	0.05	---	0.46***		
Residual Science Gain	0.04	0.00	-0.02	---	0.30***	0.44***	
Residual Social Studies Gain	0.09	-0.07	-0.03	---	0.27***	0.39***	0.31***
Panel B: Including political tolerance (N=177)							
Locus of control	0.47***						
Self-Esteem	0.34***	0.56***					
Political Tolerance	-0.04	0.17**	0.09				
Residual Math Gain	-0.07	-0.05	-0.09	0.03			
Residual ELA Gain	0.09	0.09	0.06	0.02	0.40***		
Residual Science Gain	0.10	0.07	-0.05	-0.03	0.28***	0.39***	
Residual Social Studies Gain	0.06	-0.06	-0.05	-0.11	0.21***	0.42***	0.32***

*** - $p < .01$, ** - $p < .05$, * - $p < .10$

Note. Samples restricted to students with complete responses across all measures.

Source. Authors' calculations.

In both Panel A and Panel B, the measures of students' non-cognitive skills are strongly correlated with one another, in spite of their relatively low levels of internal reliability, as are the group of achievement gains measures. On the other hand, the two groups of measures—non-cognitive skills and achievement gains—are not strongly correlated with each other. These findings contrast with the work of West et al. (2016), who found significant, but very weak, relationships between grit and achievement gains in math and ELA. Political tolerance does not appear to be related to grit, self-esteem, or achievement gains (Panel B) but is significantly and positively related to locus of control.

The raw distributions of the scores on our measures between students awarded an LSP scholarship and the control group of students who were not awarded appear in Figure 1. The figure plots kernel density estimates of the distributions for each of our four measures for students awarded and not awarded an LSP scholarship in 2012-13. While the plots do not control for student demographics and achievement, they are nevertheless informative. In particular, the similarity between the two distributions in each graph is quite striking, suggesting little average difference between the two groups. This is confirmed by Kolmogorov-Smirnov tests, which fail to reject the null of similar distributions in each case (grit: $p = 0.29$; locus of control: $p = 0.35$; self-esteem: $p = 0.41$; political tolerance: $p = 0.55$).

The results presented in Figure 1 do not suggest strong differences in non-cognitive skills and political tolerance between the two groups of students after two years of potential program participation. Nevertheless, these findings are based on simple comparisons between the two groups. Next, we examine whether the null findings presented in Figure 1 persist when controlling for observational differences between the two groups using multiple regression analysis.

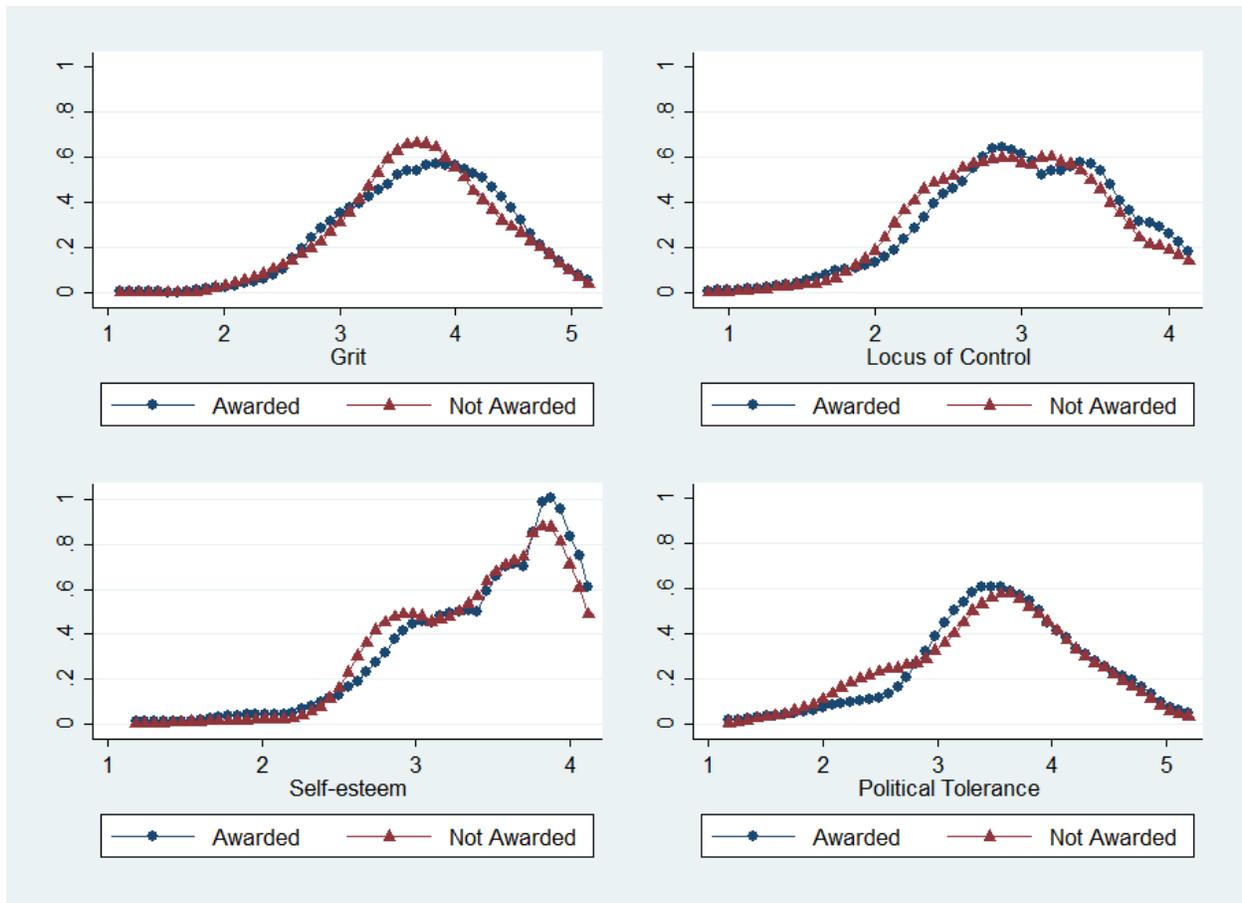


Figure 1. Kernel density estimated distributions of non-cognitive skills and tolerance measures comparing students receiving an LSP scholarship and those who did not receive a scholarship.

Tables 3 and 4 present results of regression models designed to improve model precision by controlling for various baseline characteristics. Table 3 presents results for models focusing on grit and locus of control and Table 4 presents models focusing on self-esteem and political tolerance. In both tables, columns 1 and 4 present simple models analogous to the distributional analysis presented in Figure 1. Columns 2 and 5 include controls for student demographics along with fixed effects for grade and the number of school preferences offered at application.²⁷

Columns 3 and 6 additionally include controls for student math and ELA achievement in the

²⁷ Families could offer up to 5 school preferences on their application. In order to control for unobservable differences between families offering more or fewer school preferences, we include the total number of choices offered as a vector of dummy variables.

2011-12 school year. These analyses are limited to the subset of students in our sample who took either the Louisiana LEAP or iLEAP exam in grades 3 through 7 in that year.

The results presented in Tables 3 and 4 suggest limited differences between students receiving and not receiving an LSP scholarship on all measures. Even after controlling for several baseline covariates, the general pattern of insignificant differences between the two groups suggested in Figure 1 persists. The results presented in Tables 3 and 4 are measured imprecisely, as indicated by low R-squared values. This is consistent with the low internal reliability scores reported for these scales. While overall model precision generally improves with the inclusion of additional covariates, all models perform poorly in parsing away error variance as none of the adjusted R-squared values surpass 0.09. While we expected the measures to include significant measurement error, given the lack of studies validating the included scales via phone surveys or in samples of children as young as some of those included in our sample, the results presented in Tables 3 and 4 give us little confidence in these models.

Finally, an examination of the estimated coefficients for the baseline covariates in Tables 3 and 4 reveals some interesting relationships in our sample. Females report higher levels of grit but do not differ substantially from males on the remaining measures. Moving in the last two years is associated with lower levels of grit and self-esteem but higher levels of political tolerance. Finally, student achievement has little predictive value for the set of non-cognitive skills measures; however, students with higher baseline math achievement appear to be less tolerant than other students and students with higher baseline ELA achievement appear to be relatively more tolerant.

Table 3

Regression adjusted relationships between grit and locus of control and LSP scholarship receipt

	Grit			Locus of Control		
	(1)	(2)	(3)	(4)	(5)	(6)
LSP Awarded	0.04 (0.04)	-0.02 (0.05)	0.01 (0.07)	0.06 (0.04)	0.04 (0.04)	0.04 (0.06)
Female		0.14*** (0.04)	0.06 (0.07)		0.00 (0.04)	-0.05 (0.06)
Black		0.09 (0.10)	0.31 (0.27)		0.24* (0.13)	0.41* (0.24)
White		-0.11 (0.12)	-0.14 (0.30)		0.17 (0.15)	0.15 (0.25)
Hispanic		0.27* (0.15)	0.31 (0.32)		0.33** (0.17)	0.48 (0.29)
Special Education		-0.23** (0.11)	-0.69*** (0.18)		-0.11 (0.10)	-0.16 (0.19)
Moved		-0.10** (0.05)	-0.10 (0.09)		0.03 (0.04)	0.09 (0.07)
Mother's Education						
Finished High School		0.12 (0.11)	-0.11 (0.15)		-0.03 (0.11)	-0.10 (0.10)
Went to College but Did Not Finish		0.21* (0.11)	-0.05 (0.15)		0.12 (0.11)	0.07 (0.10)
Finished College		0.18 (0.11)	-0.08 (0.15)		0.05 (0.11)	0.08 (0.10)
Standardized Math			0.05 (0.05)			0.05 (0.05)
Standardized ELA			0.02 (0.05)			0.07 (0.05)
Grade Fixed Effects		X	X		X	X
Choices Offered Fixed Effects		X	X		X	X
N	999	924	330	999	924	330
Adj. R-squared	0.00	0.04	0.06	0.00	0.04	0.09

Note. Math and ELA achievement has been standardized to the state testing distribution by grade for students taking the iLEAP or LEAP exams in grades 3 through 7 in 2011-12. "Choices Offered Fixed Effects" are indicator variables for the number of school preferences listed. Heteroskedasticity robust standard errors are presented in parentheses. *** - $p < .01$, ** - $p < .05$, * - $p < 0.10$

Source. Authors' calculations.

Table 4

Regression adjusted relationships between Self-esteem and Political Tolerance and LSP scholarship receipt

	Self-esteem			Political Tolerance		
	(1)	(2)	(3)	(4)	(5)	(6)
LSP awarded	0.03 (0.03)	-0.03 (0.04)	0.00 (0.05)	0.08 (0.09)	0.05 (0.11)	0.11 (0.13)
Female		0.02 (0.03)	0.01 (0.05)		-0.05 (0.10)	-0.12 (0.11)
Black		0.13* (0.07)	0.25* (0.13)		0.05 (0.16)	0.06 (0.25)
White		0.03 (0.09)	0.06 (0.15)		-0.19 (0.23)	-0.07 (0.31)
Hispanic		0.21** (0.09)	0.39** (0.16)		-0.06 (0.44)	-0.03 (0.47)
Special Ed		-0.15* (0.08)	-0.26** (0.11)			
Moved		-0.09** (0.04)	0.00 (0.06)		0.18* (0.11)	0.13 (0.11)
Mom's education						
Finished high school		0.00 (0.08)	0.00 (0.11)		-0.12 (0.17)	-0.17 (0.22)
Went to college but did not finish		0.07 (0.07)	0.06 (0.11)		-0.18 (0.16)	-0.23 (0.21)
Finished college		0.06 (0.07)	0.18* (0.10)		0.00 (0.15)	-0.01 (0.20)
Std. math			0.04 (0.04)			-0.25*** (0.08)
Std. ELA			-0.02 (0.04)			0.25*** (0.09)
Grade FE		X	X		X	X
Choices offered FE		X	X		X	X
N	999	924	330	247	238	211
Adj. R-squared	0.00	0.08	0.09	0.00	-0.01	0.04

Note. Math and ELA achievement has been standardized to the state testing distribution by grade for students taking the iLEAP or LEAP exams in grades 3 through 7 in 2011-12. Heteroskedasticity robust standard errors are presented in parentheses. *** - $p < .01$, ** - $p < .05$, * - $p < 0.10$

Source. Authors' calculations.

In summary, the results presented in this section do not suggest that students awarded an LSP scholarship differed from students who did not receive a scholarship on measures of non-cognitive skills and political tolerance two years later. These results hold both in sample comparisons of scholarship receipt status as well as in more complex analyses controlling for several covariates collected at baseline.²⁸

6. Conclusion

The results presented in this paper represent the first attempt to examine differences in students' self-reported measures of non-cognitive skills and political tolerance due to participation in a private school voucher program. In general, our findings do not suggest that students who were awarded an LSP scholarship differed significantly from students who did not receive a scholarship award two years after initial assignment on scales measuring individual grit, locus of control, self-esteem, and political tolerance. However, internal reliability scores indicate that several of our scales do not perform well in our samples. In particular, it is highly likely that the observed null results for both grit and locus-of-control are in part reflective of attenuation bias. We therefore do not interpret these results as conclusive and advise against making any claims about the LSP's causal impacts on student self-reports of non-cognitive skills based on these findings.

This research comes at an important time for the evaluation of educational interventions. Non-cognitive measures are increasingly being used in education evaluations. However researchers continue to recommend caution in expanding their use in policy evaluation, citing

²⁸ We conduct a robustness check of these results by restricting the sample to those who received or did not receive an LSP scholarship via lottery. We find no differences between students who did or did not receive a scholarship. Details of this experimental analysis can be found in Appendix B.

issues that may produce misleading results—such as reference-group bias—and a need to develop more accessible yet valid measures of non-cognitive skills (Dobbie & Fryer, 2015; Duckworth & Yeager, 2015; West et al. 2016).²⁹ Meanwhile, researchers are encouraged to use behavioral measures of non-cognitive skills as well as attainment benchmarks such as high school graduation and college enrollment to capture intermediate and long-term effects of educational programs on important student outcomes besides achievement, at least until better attitudinal measures of these traits can be developed and validated.

Indeed, it is important to highlight several caveats that should be taken into account when analyzing our findings. First, these findings are based on a subset of individuals volunteering to participate in the phone survey, ultimately representing 11 percent of the eligible applicants in 2012. While we present evidence indicating that survey participants do not differ strongly from the full population of LSP applicants on observable factors, we cannot rule out that our sample differs from other program participants on unobservable dimensions. Second, and perhaps more importantly, the scales used in this analysis have been validated neither for use in a phone survey nor in samples of children as young as some of those included in our survey. As we argue in the results section, these factors likely play a substantial role in the large amounts of measurement error observed in our models.

With a growing research base demonstrating the important role of non-cognitive skills in lifelong outcomes, as well as a long standing view of the important role of education in developing civic skills, it is important for future evaluations of choice programs to explore

²⁹ Both studies (Dobbie & Fryer, 2015; West et al., 2016) have provided evidence suggesting that observed differences in levels of grit among charter school students may be explained by students facing school environments with higher behavioral expectations. In both cases, reference group bias was motivated by divergent findings between the included non-cognitive skills measures and other positive outcomes.

PRELIMINARY: DO NOT CITE WITHOUT AUTHORS' PERMISSION

outcomes beyond achievement and attainment. We encourage researchers to continue this line of research and to also consider the difficulties we experienced in our own attempt to do so.

7. References

- Abdulkadiroglu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *The Quarterly Journal of Economics*, 126(2), 699–748.
- Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. (2011). *Personality psychology and economics* (NBER Working Paper No. 16822). Cambridge, MA: National Bureau of Economic Research.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Baumeister, R. F., Vohs, K. D., & Trice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, 16(6), 351–355.
- Berliner, D. C., & Biddle, B. J. (1996). *The manufactured crisis: Myths, frauds, and the attack on America's public schools*. New York, NY: Basic Books.
- Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas Ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43 (4): 972–1059.
- Bryk, A., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.
- Campbell, D. E. (2002). *The civic side of school reform: How do school vouchers affect civic education?* Princeton, NJ: Working paper of the Center for the Study of Democratic Politics.
- Candal, C. & Glenn, C. (2012) Race Relations in an Evangelical and a Catholic urban high school. *Journal of School Choice* 6(1), 82–103
- Chingos, M. M., & Peterson, P. E. (2015). Experimentally estimated impacts of school vouchers on college enrollment and degree attainment. *Journal of Public Economics*, 122, 1–12.
- Cowen, J. M., Fleming, D. J., Witte, J. F., Wolf, P. J., & Kisida, B. (2013). School vouchers and student attainment: Evidence from a state-mandated study of Milwaukee's Parental Choice Program. *Policy Studies Journal*, 41(1), 147–168.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando, FL: Harcourt Brace Jovanovich College Publishers.
- Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4), 738–782.

- Dee, T. S., & West, M. R. (2011). The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis*, 33(1), 23–46.
- Dewey, J. (1916). *Democracy and education*. New York, NY: The Macmillian Company.
- Dobbie, W., & Fryer, R. G. (2015). The medium-term impacts of high-achieving charter schools. *Journal of Political Economy*, 123(5): 985-1037.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). GRIT: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101.
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale. *Journal of Personality Assessment*, 91(2), 166–174.
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44(4), 237–251.
- Elder, T., & Jepsen, C. (2014). Are Catholic primary schools more effective than public primary schools? *Journal of Urban Economics*, 80(28-38).
- Egalite, A. J., Mills, J. N., & Greene, J. P. (2015). The softer side of learning: Measuring student' non-cognitive skills. *Improving Schools*.
- Eisenstein, M. (2006). Rethinking the relationship between religion and political tolerance in the US. *Political Behavior*, 28(4), 327–348.
- Farkas, G. (2003). Cognitive skills and non-cognitive traits and behaviors in stratification processes. *Annual Review of Sociology*, 29, 541–562.
- Figlio, D., & Ludwig, J. (2012). Sex, drugs, and Catholic schools: Private schooling and non-market adolescent behaviors. *German Economic Review*, 13(4), 385–415.
- Fleming, D. J., Mitchell, W., & McNally, M. (2014). Can markets make citizens? School vouchers, political tolerance, and civic engagement. *Journal of School Choice*, 8(2), 213-236.
- Gibson, J. (2010). The political consequences of religiosity: Does religion always cause political intolerance? In A. Wolfe & I. Katznelson (Eds.), *Religion and democracy in the United States: Danger or Opportunity?* (pp. 147–175). Princeton, NJ: Princeton University Press.
- Gutmann, A. (2003). Assessing arguments for school choice: Pluralism, parental rights, or educational results? In A. Wolfe (Ed.), *School Choice: The Moral Debate* (pp. 126–148). Princeton, NJ: Princeton University Press.

PRELIMINARY: DO NOT CITE WITHOUT AUTHORS' PERMISSION

- Hanushek, E. A., & Woessmann, L. (2009). *Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation* (NBER Working Paper No. 14633). Cambridge, MA: National Bureau of Economic Research.
- Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry*, 46(3).
- Heckman, J. J., & Kautz, T. D. (2012). *Hard evidence on soft skills* (NBER Working Paper No. 18121). Cambridge, MA: National Bureau of Economic Research.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and non-cognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411–482.
- Henig, J. R. (1999). Call for choice and radical reform. In *Rethinking School Choice: Limits of the Market Metaphor* (pp. 3–25). Princeton, NJ: Princeton University Press.
- Howell, W. G., & Peterson, P. E. (2002). School choice and American democracy. In *The Education Gap: Vouchers and Urban Schools*. Washington, DC: Brookings Institution.
- King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 191–205.
- Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of non-cognitive factors in explaining differences in educational attainment and earnings. *Social Science Research*, 37, 888–902.
- Mills, J. N. (2013). Behaviors, noncognitive skills, and school choice. *Journal of School Choice*, 7(3), 372–388.
- Mills, J. N. & Wolf, P. J. (2016). The effects of the Louisiana Scholarship Program on Student Achievement after two years (Louisiana Scholarship Program Evaluation Report #1). Fayetteville, AR: School Choice Demonstration Project, University of Arkansas.
- Mocan, H. N., & Tekin, E. (2007). *Catholic schools and bad behavior: A propensity score matching analysis* (NBER Working Paper No. 9172). Cambridge, MA: National Bureau of Economic Research.
- Naumann, L. P., & John, O. P. (2013). Toward a domain-specific approach to cultural differences: The influence of cultural values and reference-group standards on self-reported personality. *Unpublished Manuscript*.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

- Rotter, J. R. (1966). Generalized expectations for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1), 1–28.
- Scanlan, M. (2008). The grammar of Catholic schooling and radically” Catholic” Schools. *Catholic Education: A Journal of Inquiry and Practice*, 12(1), 25–54
- Sullivan, J., Piereson, J., & Marcus, G. (1982). *Political tolerance and American democracy*. Chicago, IL: University of Chicago Press.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72(2), 271–322.
- Tuttle, C. C., Gill, B., Gleason, P. M., Knechtel, V., Nichols-Barrer, I., & Resch, A. (2013). *KIPP Middle Schools: Impacts on Achievement and Other Outcomes, Final Report*. Washington, DC: Mathematica Policy Research.
- United States Department of Education. (1986). *High School and Beyond, 1980: Sophomore and Senior Cohort Third Follow-up*. Washington, DC: National Center for Education Statistics.
- Waddington, R. J., & Berends, M. (2015). *Vouchers in the crossroads: Heterogeneous impacts on student achievement and attendance across private schools in Indiana*. (Paper presented at the annual meeting of the Association for Public Policy Analysis and Management). Miami, FL.
- Warren, J. R. (2011). Graduation rates for choice and public school students in Milwaukee, 2003-2009. *School Choice Wisconsin*.
- West, M. R., Kraft, M. K., Finn, A. S., Martin, R., Duckworth, A. L., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2016). Promise and paradox: Measuring students’ non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*, 38(1): 148-170.
- Wolf, P. J. (2005). School choice and civic values. In *Getting Choice Right* (pp. 210–244). Washington, DC: Brookings Institution Press.
- Wolf, P. J. (2008). School voucher programs: What the research says about parental school choice. *Brigham Young University Law Review*, 2008:2, 415-46.
- Wolf, P. J., Kisida, B., Gutmann, B., Puma, M., Eissa, N., & Rizzo, L. (2013). School vouchers and student outcomes: Experimental evidence from Washington, DC. *Journal of Policy Analysis and Management*, 32(2), 246–270.
- Wolf, P. J., Peterson, P. E., & West, M. R. (2001). *Results of a school voucher experiment: The case of Washington, DC after two years* (Working Paper No. 01-05). Cambridge, MA: Program on Education Policy and Governance, Harvard University.

PRELIMINARY: DO NOT CITE WITHOUT AUTHORS' PERMISSION

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: Massachusetts Institute of Technology.

8. Appendix A

The LSP Scholarship Award Algorithm

When the LSP was expanded to a statewide program in 2012, the Louisiana Department of Education also changed the lottery process determining scholarship awards. While the original application process in the New Orleans pilot version of the LSP limited families to submitting the name of only one private school for admission, the revised application process allowed individuals to offer up to five private school preferences. This new lottery process is similar to the deferred acceptance lotteries used in New York City to assign students to schools through the city's public school choice program (see Abdulkadiroglu, Pathak, & Roth, 2005). The deferred acceptance algorithm is designed to encourage families to reveal their true school preference rankings and thereby reduce the likelihood of gaming.

While it is not the case that all eligible LSP applicants were awarded scholarships through a lottery process in the 2012-13 school year, we can isolate cases in which lotteries occurred in order to perform an experimental evaluation of the program.

Specifically, eligible LSP applicants are allowed to submit up to five private school preferences and the LSP lottery algorithm attempts to place students into schools while taking into account several lottery priorities. First, students with disabilities and "multiple birth siblings"³⁰ are manually awarded LSP scholarships if there is available space at their given school preference. Remaining students are grouped into one of six priority categories:

- **Priority 1** - Students who received LSP scholarships in the prior school year who are applying to the same school
- **Priority 2** - Siblings of Priority 1 awardees in the current round

³⁰ "Multiple birth siblings" are twins, triplets, etc.

- **Priority 3** - Students who received LSP scholarships in the prior school year who are applying to a different school
- **Priority 4** – New applicants who attended public schools that received a “D” or “F” grade in Louisiana’s school accountability system at baseline
- **Priority 5** – New applicants who attended public schools that received a “C” grade in Louisiana’s school accountability system at baseline
- **Priority 6** – New applicants who are applying for kindergarten placements

The first stage of the LSP award process is summarized in Figure 1. The process begins by attempting to place all Priority 1 category students into their first choice school.³¹ The algorithm first groups Priority 1 students applying to the same school and grade combination and then checks the number of available seats for that grouping. If there are more seats than applicants, all students receive an LSP scholarship. If there are no seats available, no students in the given group receive a scholarship. Finally, if there are more applicants than seats, students are awarded LSP scholarships through a lottery. Once the process is complete for all Priority 1 students, the algorithm attempts to place Priority 2 students into their first choice school. After cycling through all remaining priority categories, the LSP algorithm moves to the second stage of the allocation process by attempting to place students who have yet to receive a scholarship in their second choice schools. The LSP algorithm continues until all eligible applicants have either been awarded or not awarded an LSP scholarship.

³¹ By definition, the first choice school for a priority 1 category student is the school they previously attended in the New Orleans pilot version of the program.

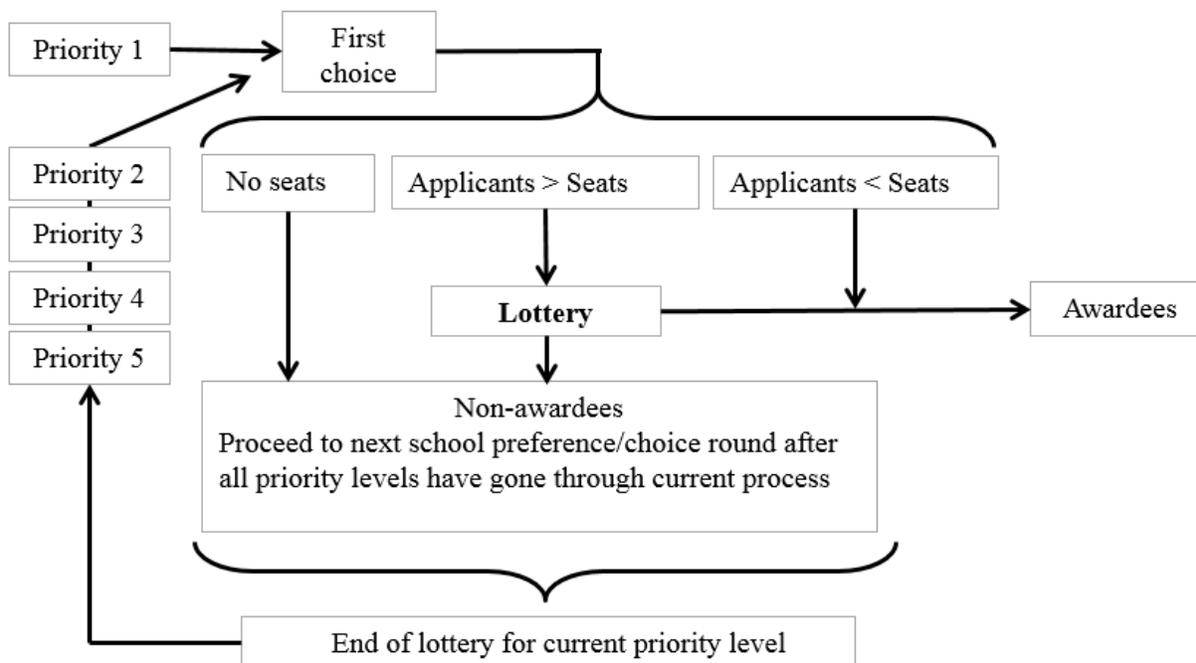


Figure 1. First stage of the Louisiana Scholarship Program award allocation process for the 2012-13 school year. This figure illustrates the iterative process used to allocate LSP scholarships to students. In addition, this figure highlights the fact that only a subset of students was awarded LSP scholarships via lotteries. Our analysis focuses on isolating lotteries for one’s first choice school.

Only a subset of eligible applicants were awarded or not awarded an LSP scholarship via a lottery process. Specifically, only those students in priority categories one through six whose school-grade combination had more applicants than available seats participated in a lottery. Fortunately, using data on student characteristics and school preferences, we can identify the subset of eligible applicants who experienced a lottery process.³² The subset of eligible LSP applicants whose scholarship outcome was determined by a lottery are focus of the following

³²We identify a lottery as occurring when the percentage of students awarded an LSP scholarship falls between 0 and 100 percent for a given school preference by grade by priority category combination. For example, if 60 percent of Priority 1 category students applying to third grade at school “A” as their first choice school actually received scholarships, we identify all students in that combination as having been subject to a lottery.

robustness check of this paper's results (Appendix B) and a separate study in this evaluation estimating the impact of LSP participation on student achievement (Mills & Wolf, 2016).³³

³³ We will focus on this subset of lottery participants to estimate the effects of the LSP on student achievement after two years of program participation because these are the only applicants for whom LSP scholarship award was randomly determined. This focus on oversubscription lotteries suggests our analysis may be capturing the most favorable estimates of the program's effectiveness, as higher quality schools are often more likely to be oversubscribed than lower quality schools (Abdulkadiroglu, Angrist, Dynarski, Kane, & Pathak, 2011).

9. Appendix B

Robustness Check: Experimental Analysis

As a robustness check, we examine if the null results identified in the descriptive analyses presented in the preceding section hold among a subsample of phone survey participants whose LSP scholarship award was determined by a lottery. LSP scholarships were awarded to students through a matching algorithm designed to take into account student school preferences as well as a set of priorities established by the Louisiana Department of Education (LDE). While all students were subject to the matching algorithm, LSP scholarships were only awarded by lottery in cases when there were more students applying in the same priority category than seats available to the same grade in the same school.³⁴ Thus, as a check of the robustness of the null results presented in Tables 3 and 4, we examine the extent to which these findings persist in the subsample of eligible applicants participating in binding lotteries. By focusing on lotteries, we will be limiting our sample; however we will be providing a better control for unmeasurable factors driving selection into private schooling. While we should not expect to find substantially different results in this group, such a finding would raise concerns regarding our primary analyses.

Our focus on binding lotteries requires a change in the model used to estimate differences in students receiving and not receiving an LSP scholarship. In particular, we employ a two-stage least squares (2SLS) model which allows an estimation of the effect of LSP usage (also known as the Local Average Treatment Effect) on the non-cognitive skill and tolerance measures of interest. This process first requires predicting the likelihood that a student enrolls in a private school using their LSP scholarship lottery outcome as a predictor along with a series of controls

³⁴ See Appendix A for a more detailed description of the LSP scholarship award algorithm.

for demographics, baseline achievement, and individual risk set. This predicted usage variable is then substituted for observed usage in a model predicting one of the given dependent variables:

grit, locus of control, self-esteem, and political tolerance. The specific 2SLS model is:

$$1. \quad E_i = \sum \pi_j R_{ji} + \delta T_i + \mathbf{X}\boldsymbol{\beta} + \mathbf{A}\boldsymbol{\rho} + u_i$$

$$2. \quad M_i = \sum \alpha_j R_{ji} + \tau \hat{E}_i + \mathbf{X}\boldsymbol{\gamma} + \mathbf{A}\boldsymbol{\theta} + \epsilon_i$$

Where:

- R is a fixed effect for a student's first choice school lottery or "risk set"³⁵
- E is a variable indicating if a student used an LSP scholarship to enroll in a private school
- T is a variable indicating if a student received an LSP scholarship to their first choice school
- M is one of the four outcome measures of interest in this study: Grit, Locus of Control, Self-esteem, and Political Tolerance
- X is a vector of student demographics³⁶
- A is a vector of variables capturing student achievement in 2011-12³⁷

By including fixed effects for binding lotteries, we are able to ensure that we are comparing individuals whose scholarship allocation result was determined randomly.

Nevertheless, given our 10% response rate, it is true that we rarely recover all students involved in a binding lotteries. The lottery fixed effects strategy effectively uses the outcomes of observed scholarship recipients and non-recipients to stand in for the other students in the lottery who did

³⁵ We use standard errors that account for clustering of students within their binding lotteries to avoid the potential for biased inference (Angrist & Pischke, 2009).

³⁶ Demographic controls include gender, race/ethnicity, an indicator of student mobility, mother's education, and variables capturing the number of school preferences offered at application.

³⁷ Regressions including student achievement are restricted to students who took the iLEAP or LEAP exam in math and ELA in grades 3 through 7 in 2011-12.

not participate in the phone survey. Admittedly, this is a drawback of our analysis; and we therefore caution the reader to take this design feature into account when interpreting our results.

Table B1 presents the results of the 2SLS estimations of the differences between LSP scholarship users and other students on our non-cognitive skills and political tolerance measures. Column 1 presents results for simple models that only include risk set fixed effects; column 2 provides the results from specifications that additionally control for student demographics; and column 3 presents results for models additionally controlling for student baseline achievement. The latter models are restricted to the subset of students who took the Louisiana state assessments in grades 3 through 7 in 2011-12. Across all models, the results from first stage regressions suggest winning an LSP scholarship is highly predictive of use: LSP winners, on average, are over 85 percentage points more likely to enroll in a private school across all models; and the reported joint-F statistics meet Staiger and Stock's (1997) recommended threshold of 10.

The results presented in Table B1 do not generally suggest students using an LSP scholarship to their first choice school differ from other students on the self-reported measures of non-cognitive skills and political tolerance in nearly every model. The lone exception is that we find that LSP scholarship users on average report significantly higher scores on the Locus of Control Scale in a model accounting for lotteries, student demographics, and student baseline achievement scores. This finding is somewhat surprising, given the overall insignificant results observed for the companion model presented in Table 4. Yet given a low degree of internal consistency across items in the Locus of Control Scale ($\alpha = 0.47$) we are uncertain what underlying trait we have truly captured in this measure.

Table B1

Student-level relations between self-reported measures of non-cognitive skills and tolerance and LSP scholarship receipt in binding lotteries

Dependent Variable	without covariates (1)	+ demographic controls (2)	+ student achievement (3)
Grit	0.05 (0.09)	-0.02 (0.09)	0.12 (0.16)
Locus of Control	0.04 (0.08)	0.03 (0.08)	0.18* (0.11)
Self-esteem	-0.06 (0.07)	-0.09 (0.08)	-0.02 (0.12)
Model summary			
N	639	587	202
Risk sets	280	264	105
First stage joint F	974.3	78.6	23.8
Political tolerance	0.00 (0.20)	-0.08 (0.23)	-0.03 (0.27)
N	157	150	136
Risk sets	83	79	76
First stage joint F	229.3	17.9	16.4

*** - $p < .01$, ** - $p < .05$, * - $p < 0.10$

Note. All models include risk set fixed effects. Across all models, winning an LSP scholarship is highly predictive of use: all estimated coefficients on LSP awarded are over .85 and have reported p-values of less than .001.

Source. Authors' calculations

There are at least two reasons for this discrepancy. First, it is important to note that the focus on binding lotteries, in addition to the requirement of baseline achievement data in the model in question, restricts the sample on which this result is based (sample size of 202 compared to a sample size of 330 in the primary analysis). In addition, the results presented in Table 5 are based on local average treatment effects—or the estimated effects of LSP usage for those students whose treatment assignment influences their take up. In contrast, the observational analyses presented in Table 3 estimate relationships based on the mere scholarship award outcome. Nevertheless, while these points may explain the significant finding for locus of

PRELIMINARY: DO NOT CITE WITHOUT AUTHORS' PERMISSION

control, the generally insignificant results presented in Table 5 generally corroborate the insignificant findings presented in Tables 3 and 4.

About the Authors

Jonathan N. Mills is a Postdoctoral Fellow at the Education Research Alliance for New Orleans at Tulane University. His research focuses on the effects of school choice programs on student achievement and non-academic outcomes, as well as the benefits and unintended consequences of college financial aid programs. Mills received his Ph.D. in education policy from the University of Arkansas in 2015. He additionally holds a Bachelor of Science and a Master of Arts in economics from the University of Missouri.

Albert Cheng is a doctoral candidate in education policy in the Department of Education Reform at the University of Arkansas. He has published in journals such as the *Economics of Education Review* and *Social Science Quarterly* on non-cognitive skills, civic values, and school choice. He received his B.A. in mathematics from the University of California, Berkeley and was a public high school math teacher in the San Francisco Bay Area prior to joining the University of Arkansas.

Collin Hitt is a doctoral candidate in education policy in the Department of Education Reform at the University of Arkansas, where he is also a research fellow for *Charassein: The Character Assessment Initiative*. His research explores new methods of measuring character skills, and has appeared in *Economics of Education Review* and *Education Next*. He has also authored numerous studies of school choice. He holds a B.A. in philosophy and political science from Southern Illinois University.

Patrick J. Wolf is Distinguished Professor of Education Policy and 21st Century Endowed Chair in School Choice at the University of Arkansas in Fayetteville. He has authored, co-authored, or co-edited four books and over 100 journal articles, book chapters, and policy reports on school choice, civic values, public management, special education, and campaign finance. He received his Ph.D. in Political Science from Harvard University in 1995.

Jay P. Greene is Distinguished Professor of Education Policy, 21st Century Endowed Chair in Education Reform, and Head of the Department of Education Reform at the University of Arkansas. His current areas of research interest include school choice, culturally enriching field trips, and the effect of schools on non-cognitive skills and civic values. He received his B.A. in history from Tufts University in 1988 and his Ph.D. in Political Science from Harvard University in 1995.

About the SCDP

Housed within the Department of Education Reform at the University of Arkansas, the School Choice Demonstration Project (SCDP) is an education research center dedicated to the non-partisan study of the effects of school choice policy. Led by Dr. Patrick J. Wolf, the SCDP's national team of researchers, institutional research partners and staff are devoted to the rigorous evaluation of school choice programs and other school improvement efforts across the country. The SCDP is committed to raising and advancing the public's understanding of the strengths and limitations of school choice policies and programs by conducting comprehensive research on what happens to students, families, schools and communities when more parents are allowed to

PRELIMINARY: DO NOT CITE WITHOUT AUTHORS' PERMISSION

choose their child's school. Reports from past SCDP studies are available at <http://www.uaedreform.org/school-choice-demonstration-project/>.

About ERA-NOLA

The Education Research Alliance for New Orleans is based within the School of Liberal Arts at Tulane University. The objective of ERA-New Orleans is to provide objective, rigorous, and useful research to improve schools in New Orleans and beyond. For more information about ERA-New Orleans, including some of the studies mentioned in this brief, please visit www.educationresearchalliancenola.com.