The MPCP Longitudinal Educational Growth Study

Third Year Report

John F. Witte
University of Wisconsin–Madison

Joshua M. Cowen
University of Kentucky

David J. Fleming
Furman University

Patrick J. Wolf
University of Arkansas

Meghan R. Condon
University of Wisconsin–Madison

Juanita Lucas-McLean
Westat

SCDP Milwaukee Evaluation Report #15

April 2010

The University of Arkansas was founded in 1871 as the flagship institution of higher education for the state of Arkansas. Established as a land grant university, its mandate was threefold: to teach students, conduct research, and perform service and outreach.

The College of Education and Health Professions established the Department of Education Reform in 2005. The department's mission is to advance education and economic development by focusing on the improvement of academic achievement in elementary and secondary schools. It conducts research and demonstration projects in five primary areas of reform: teacher quality, leadership, policy, accountability, and school choice.

The School Choice Demonstration Project (SCDP), based within the Department of Education Reform, is an education research center devoted to the non-partisan study of the effects of school choice policy and is staffed by leading school choice researchers and scholars. Led by Dr. Patrick J. Wolf, Professor of Education Reform and Endowed 21st Century Chair in School Choice, 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 choose their child's school.

The MPCP Longitudinal Educational Growth Study

Third Year Report

John F. Witte
University of Wisconsin–Madison

Joshua M. Cowen
University of Kentucky

David J. Fleming
Furman University

Patrick J. Wolf
University of Arkansas

Meghan R. Condon
University of Wisconsin–Madison

Juanita Lucas-McLean

SCDP Milwaukee Evaluation
Report #15

April 2010

SCHOOL CHOICE
DEMONSTRATION PROJECT

School Choice Demonstration Project
Department of Education Reform
University of Arkansas
201 Graduate Education Building
Fayetteville, AR 72701

479-575-6345

http://www.uark.edu/ua/der/SCDP/Research.html

CONTENTS:

EXECUTIVE SUMMARY

This is the third-year report in a five-year evaluation of the Milwaukee Parental Choice Program (MPCP). The MPCP, which began in 1990, provides government-funded vouchers for low-income children to attend private schools in the City of Milwaukee. The maximum voucher amount in 2008-09 was $6,607, and approximately 20,000 children used a voucher to attend either secular or religious private schools. The MPCP is the oldest and largest urban school voucher program in the United States. This evaluation was authorized by Wisconsin Act 125 enacted in 2005.

The general purposes of the evaluation are to analyze the effectiveness of the MPCP in terms of longitudinal student achievement growth and educational attainment as measured by high school graduation rates. The former will be primarily accomplished by measuring and estimating student growth in achievement as measured by the Wisconsin Knowledge and Concepts Examinations (WKCE) in math and reading in 2006-07, grades 3 through 8 over a five-year period. The latter will be accomplished by following the 2006-07 8th and 9th grade cohorts over a five-year period or longer. The general research design consists of a comparison between a random sample of MPCP students and a matched sample of Milwaukee Public School (MPS) students.

The February 2008 baseline report (Witte, Wolf, Cowen, Fleming, & Lucas-McLean, 2008) presented sample means and standard deviations of student test scores in math and reading subjects on the November 2006 WKCE tests. The second year report, released in 2009, estimated achievement differences for one year of growth in MPCP and MPS samples. The conclusion showed no differences between the two samples of students. In this third year report, we present results from the November 2008 WKCE tests as second year student achievement growth in MPCP relative to the matched MPS sample. We provide varying descriptive statistics comparing test score means and distributions for math and reading for 2006-07 (baseline year) and 2008-09 (second outcome year) for each sample. We also analyze achievement growth using several multivariate techniques and models. The primary finding in all these comparisons is that, in general, there are few statistically significant differences between levels of MPCP and MPS student achievement growth in either math or reading two years after they were carefully matched to each other. In one of the ways of estimating these results, focusing only on those students who have remained in the public or private sector for all three years, private, voucher students are slightly behind MPS students in mathematics achievement growth. We offer several cautions in interpreting this result against the overwhelming set of results that indicate no difference in achievement growth.

We also provide in Appendix B an updated assessment of missing cases, defined as students we could not locate in the third year of the study. For our achievement analysis, of students we have tracked over the course of the study, 14 percent of the total samples drawn in 2006-07 and 2007-08 could not be located, with nearly identical percentages of missing cases coming from both MPCP and MPS. This number is considerably below our initial assumption of 20 percent per year when we conceived sample sizes and the need to refresh the samples.[1]

1 Including baseline 9th graders that are not included in the achievement study, these figures are higher: 16 percent for MPS and 28 percent for MPCP. This is not surprising, given the mobility of older students between options that include dropping out of school entirely. As part of the future attainment study, we are initiating extensive tracking efforts to determine the location of these students.

Students who could not be located, on average, had slightly higher test scores and were more likely to be white than students who remained in the study. However, in examining missing students, there were almost no differences in student characteristics between those missing from the MPCP or the MPS panels. Thus we are confident that attrition has not yet affected the relative estimates of achievement growth between sectors.

Throughout the report, we describe a range of cautions and caveats; the most important being that this is the third year of a five-year study and that student achievement trajectories often take time to change. Thus, while at the present time, in terms of achievement as measured by two years of achievement growth, we conclude that in general there is little significant difference overall between MPS students and MPCP students, this result may change in future analyses.

This report and its companion reports continue a series of annual reports on the Milwaukee Parental Choice Program conducted by the School Choice Demonstration Project (SCDP). An initial draft of this report was greatly improved based on comments from the SCDP Research Advisory Board and research team, particularly David Figlio of Northwestern University and Paul Peterson of Harvard University. All remaining errors are the responsibility of the authors alone.

This ongoing research project is being funded by a diverse set of philanthropies including the Annie E. Casey, Joyce, Kern Family, Lynde and Harry Bradley, Robertson, and Walton Family Foundations. We thank them for their generous support and acknowledge that the actual content of this report is solely the responsibility of the authors and does not necessarily reflect any official positions of the various funding organizations, the University of Wisconsin, the University of Kentucky, Furman University, the University of Arkansas, or Westat, Inc. We also express our deep gratitude to MPS, the private schools in the MPCP, and the state Department of Public Instruction for willing cooperation, advice, and assistance.

INTRODUCTION

This is the third-year report in a five-year evaluation of the Milwaukee Parental Choice Program (MPCP). This program, which began in 1990, provides government-funded vouchers for low-income children to attend private schools in the City of Milwaukee. The maximum voucher amount in 2008-09 was $6,607, and approximately 20,000 children now use a voucher to attend either secular or religious private schools. The MPCP is the oldest and largest urban school voucher program in the United States. This evaluation was authorized by Wisconsin Act 125 enacted in 2005.

The general purposes of the evaluation are to analyze the effectiveness of the MPCP in terms of longitudinal student achievement growth as measured by standardized tests, and educational attainment as measured by high school graduation rates. The former will be based on estimating student achievement growth measured by the Wisconsin Knowledge and Concepts Examinations (WKCE) in math and reading in grades baseline 3 through 8 over a five-year period. The latter will be accomplished by following the 2006-07 8th and 9th grade cohorts over a five-year period or longer.[2] The general research design consists of a comparison between a random sample of MPCP students and a matched sample of Milwaukee Public School (MPS) students. The procedures for obtaining that sample are briefly discussed in the next section and described in detail in Appendix B of Witte, Wolf, Cowen, Fleming, and Lucas-McLean (2009).

In the first year report (Witte et al., 2008), we described baseline test scores in a number of ways. The results revealed, by design, very similar baseline scores for the MPCP and matched MPS samples on the WKCE math and reading tests. The similarity was one indicator of the success of our matching algorithm. Our second year report provided one-year growth estimates from the fall of 2006 to the fall of 2007. The essence of that report was that students in private schools utilizing vouchers achieved at the same rate in math and reading as the students in the matched-MPS sample (Witte et al., 2009).

In this report we present data on two-year growth in student achievement between the fall of 2006 and the fall of 2008. We caution readers that these data are far from conclusive in terms of the final outcomes of this evaluation. Evaluations now routinely cover a number of years, not just one, especially if growth in achievement from a common baseline is the focus of the evaluation. These "value-added" models are being used to evaluate school and even classroom success in affecting student achievement. Thus, we hope to emulate and improve upon these studies over the next two years.

To begin our evaluation of achievement differences between the two samples, we first provide a range of descriptive statistics in terms of achievement growth. These include measures of central tendency, such as average gains by grade, and comparisons of the entire distribution of scores using frequency graphs. We also use a simple but intuitively appealing method to describe the chances that MPCP students did better than MPS

2 Because there is no current research on student attainment in voucher studies in the United States, we hope to be able to obtain further research funding to continue tracking these students beyond high school. This is extremely important because substantial research indicates the great advantages that accrue to students who graduate from high school, experience post-secondary education, and attain a college degree.

students in the prior two years. This method of ordinal data analysis compares the growth of each student in the MPCP sample to that of each student in the MPS group, and results in the calculation of a Somers' d statistic, a nonparametric measure that can be considered the difference in the probability that a student from one sector will have gained more than a student in the other sector from baseline to the third year.[3]

More elaborate comparisons are made using multivariate methods in which we control for the original test score of a student in 2006-07 and a number of other demographic and independent variables. Our objective is to determine if the coefficient for the variable indicating which sector the student was in at baseline (MPCP or MPS) is significantly different from zero in the statistical sense, thereby allowing us to reject the "null hypothesis" of zero difference in gains across the two school sectors.

Student mobility is a problem for all student longitudinal studies, but even more so for those conducted in high poverty areas. Mobility occurs between schools, between school districts, and through dropping out of school altogether. Mobility poses several problems and raises a number of issues. First, either dropping out of school or moving to another school district, in Wisconsin or in another state, effectively ends the acquisition of test and other data for a student. This study attrition reduces sample sizes and could introduce biased results if the missing cases are dissimilar on relevant variables depending on whether they are missing from the MPCP or the MPS panel. We examine this issue in Appendix B for the first three years, and find few systematic differences in terms of observable student characteristics in the MPS and MPCP samples. Nevertheless, we correct for differences in our results by weighting the data based on the probability of remaining in the study.

Finally, because this is not a controlled experiment, some students in our panels switch from the public to the private sector or vice versa. Although we can capture these sector switchers and test them, one important research issue is the way we account for them in the long term. Should, for example, a student who begins in the MPCP sample, but after several years moves to a public school, be counted for all the years as an MPCP student? That is what is done in most medical or drug clinical trials, and that is the method we employ in our first multivariate analysis. Another way to account for that student who switched school sectors would be to simply drop the student from the analysis once the move occurs and only estimate achievement growth for those years for which the student was in their "assigned" sector, public or private. We provide a variant of that approach as an alternative analysis by estimating achievement growth for only those students who stay in the same sector for all three years. In subsequent analyses we will present new results that match students based on the probability of sector switching. A companion report to this study (Cowen, Fleming, Witte, & Wolf, 2010) analyzes the characteristics of student switchers in greater detail.

The report to follow has three basic sections. The first analyzes achievement gains from 2006 to 2008; the second, offers some caveats and cautions; and the last offers a summary and a set of current conclusions. We analyze the sample attrition and describe our ongoing efforts to locate missing students in Appendix B. Appendix A provides descriptive statistics for variables used in our multivariate analyses.

3 This method is explicitly suggested in Ballou (2009) to compare differences between students with different teachers rather than different school sectors.

STUDENT ACHIEVEMENT GAINS: 2006 to 2008

The February 2008 baseline report (Witte et al., 2008) presented sample means and standard deviations of student test scores in math and reading subjects on the November 2006 WKCE tests. We intended these statistics to provide benchmark measures of achievement current to the onset of the longitudinal study, and to serve as indicators for the success of our sample selection methodology. In this Year 3 report, we present results from the November 2008 WKCE tests as measures of student achievement growth in MPCP relative to a matched-MPS sample over a two year growth period.

Average Math and Reading Achievement and Growth

The baseline report detailed the sample selection methodology that provides valid comparisons of MPS and MPCP students. In brief, we used students' neighborhood location, baseline test scores, and demographic information to construct the MPS sample that matched the randomly selected MPCP sample. We showed in the first year report (Witte et al., 2008) that after matching, the MPS and MPCP samples were demonstrably similar in terms of baseline test scores and other observable characteristics. This similarity was by design. Importantly, we argued that the matching algorithm—in particular the emphasis on neighborhood location—accounts for unobserved characteristics that may bias comparisons of student outcomes between the two sectors. We supported this assertion in part through rich survey data collected after the matching process, which showed very similar patterns of home environment, parental education, and educational experiences for students and their parents from the same neighborhoods, regardless of whether the students were in the MPCP or the MPS (Witte et al., 2008). Because we are confident that our matching process largely eliminated differences between the samples on factors systematically influencing student achievement, we believe that simple comparisons of Year 3 mean achievement between the sectors is a valid statistical indication of any outcome differences by the fall of 2008 between students learning in the MPS and MPCP sectors. Tables 1 and 2 provide weighted mean growth in scale scores for both the 2006-2007 (baseline) and 2008-2009 (second year outcome) academic years for students in grades 4-8.[4,5] The sample includes students who were in grades 3-6 and 8 at baseline for whom we have three years of WKCE achievement scores at this point in the study as well as students in 3rd grade in 2007-08 who were added to refresh the baseline sample and for whom we have separate baseline (fall 2007) and one year of growth (fall 2008) scores.

Because of variations in grade-level ranges in scale scores that are purposely built into the test design, comparing average group-level scale scores across grades is not appropriate. For example, we cannot simply say that MPCP 5th graders are doing better than MPS 4th graders simply because the mean is higher for 5th graders. Fifth

4 Scale scores are scores generated from basic data on the number of correct answers on a multiple choice (or other) standardized test. They fall within ranges for each grade that increase in each higher grade as tests become more complex (and the variance between students increases). They are approximately normally distributed and are integer-level measures. They are designed to measure the development of a child in each subject area and are calculated using a psychometric process called Item Response Theory or IRT.

5 Weights were created to adjust for missing test scores. Accepted research protocols call for use of weighted data in this research design.

grade achievement is measured on a separate scale from 4th grade achievement. The comparison must be relative: either limited to students within the same grade, or a measure of the extent to which students in various grades gained in achievement between the same number of years. The important point, however, is that the range of possible scores for each grade is the same for MPS and MPCP, so cross-sector comparisons within grades are valid. Tables 1 and 2 display achievement growth across one year (comparing the November 2008 test to the November 2007 test) and two years (comparing the 2008 test to the baseline 2006 test) of learning. Differences between the two groups that favor the MPCP students appear as positive numbers while differences that favor the MPS students appear as negative numbers. Tables 1 and 2 indicate a few systematic differences between the two sectors using difference-in-means tests. There is a significant difference in one- and two-year math growth scores favoring the MPCP students who were in grade 7 in 2008. There is also some evidence that MPS 8th graders had larger one year growth in math than did their MPCP counterparts. On the reading exam (Table 2), there are no significant sector effects for two year growth in reading achievement in any grade. However, MPS 5th graders exhibited larger one year growth in reading scores than did MPCP 5th graders.

Table 1: Mean Math Achievement by Grade, 2006-07 to 2008-09

Grade 2008	Group	One-Year Change (07-08)		Two-Year Change (06-08)	
		Mean Growth	s.e. (diff)	Mean Growth	s.e. (diff)
4	MPS Matched	40.2			
	MPCP	39.3			
	(Difference)	**-0.9**	**3.3**		
5	MPS Matched	24.2		59.9	
	MPCP	22.1		60.1	
	(Difference)	**-2.1**	**3.4**	**0.2**	**3.48**
6	MPS Matched	25.3		42.8	
	MPCP	20.2		43.0	
	(Difference)	**-5.1**	**3.1**	**0.2**	**3.2**
7	MPS Matched	23.2		42.5	
	MPCP	30.4		50.2	
	(Difference)	**7.2****	**3.12**	**7.7****	**3.3**
8	MPS Matched	16.3		32.2	
	MPCP	9.2		33.2	
	(Difference)	**-7.1****	**3.5**	**0.99**	**3.8**
10	MPS Matched			14.0	
	MPCP			19.2	
	(Difference)			**5.2**	**4.5**

Stars indicate MPS different from MPCP statistics at ***p<0.01, **p<0.05, *p<0.10, based on a two-tailed T-Test. Figures include only students with valid test scores in both 2006-07 and 2008-09 academic years. Mean changes may not sum perfectly due to rounding. Response weights were used for those in grades 4-8 and 10 in 2008.

Table 2: Mean Reading Achievement by Grade, 2006-07 to 2008-09

Grade 2008	Group	One-Year Change (07-08)		Two-Year Change (06-08)	
		Mean Growth	s.e. (diff)	Mean Growth	s.e. (diff)
4	MPS Matched	6.7			
	MPCP	23.1			
	(Difference)	**16.3**	**11.3**		
5	MPS Matched	8.5		13.8	
	MPCP	2.6		16.7	
	(Difference)	**-6.0****	**3.0**	**2.9**	**3.1**
6	MPS Matched	9.7		17.3	
	MPCP	15.2		20.1	
	(Difference)	**5.5**	**3.8**	**2.9**	**3.7**
7	MPS Matched	17.8		30.4	
	MPCP	16.4		32.9	
	(Difference)	**-1.4**	**3.5**	**2.5**	**3.4**
8	MPS Matched	15.6		24.5	
	MPCP	14.1		25.8	
	(Difference)	**-1.5**	**3.4**	**1.3**	**3.6**
10	MPS Matched			0.7	
	MPCP			-3.9	
	(Difference)			**-4.6**	**5.7**

Stars indicate MPS different from MPCP statistics at ***p<0.01, **p<0.05, *p<0.10, based on a two-tailed T-Test. Figures include only students with valid test scores in both 2006-07 and 2008-09 academic years. Mean changes may not sum perfectly due to rounding. Response weights were used for those in grades 4-8 and 10 in 2008.

The Distribution of Math and Reading Growth

It is possible that similar mean achievement levels, or changes in those levels, could mask differences at different levels of achievement. For example, high achieving MPS students could outperform their matched MPCP counterparts, while the opposite takes place at the bottom of the achievement distribution. In computing the means, these would cancel each other out for no effect. We demonstrate that is not the case graphically in Figures 1 and 2. The figures are Kernel densities, which are similar to histograms and represent estimates of the underlying probability distributions of the two-year change scores reported in the second to last columns of Tables 1 and 2. These figures provide perhaps the most concise comparisons of academic achievement between matched samples of MPS and MPCP students currently available. They indicate that mean growth is not only

very similar between the sectors at this point in our study, but is also distributed in much the same way. In other words, similar frequencies of MPCP and MPS students were among the highest and lowest observed growth scores.

Figure 1: November 2006-08 Math Growth for All Students in Grades 5-8 and 10

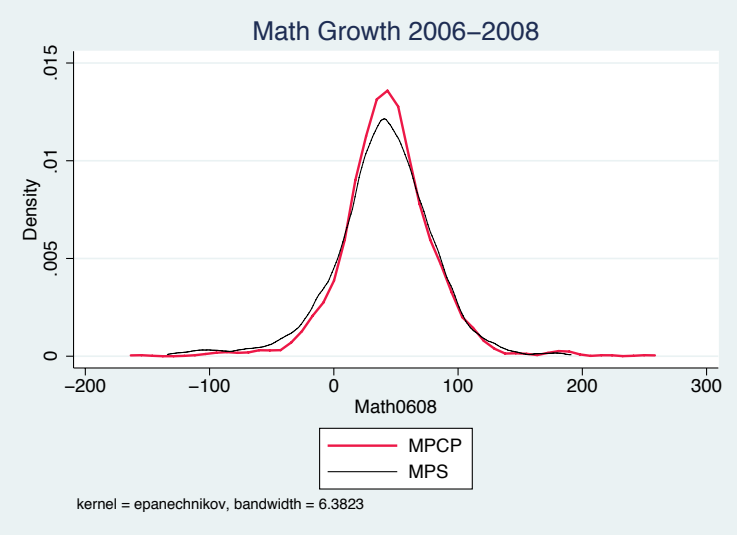
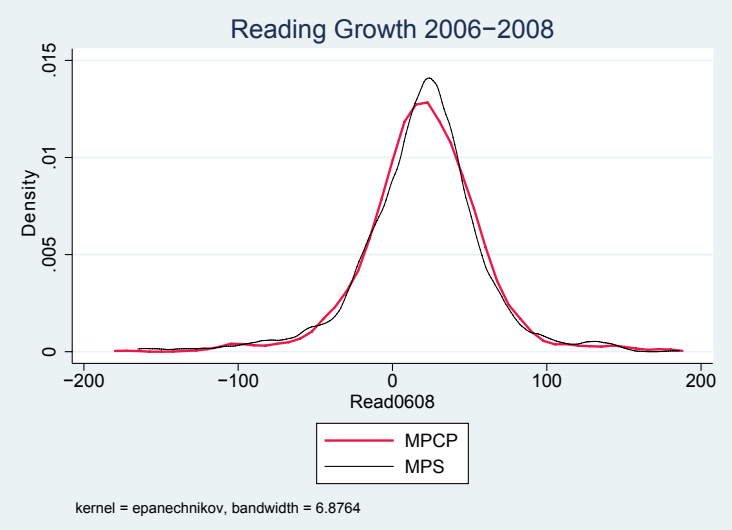


Figure 2: November 2006-08 Reading Growth for All Students in Grades 5-8 and 10



We use a method of ordinal data analysis to develop this point further. This method compares the gain score (by subject) for each MPCP student in a given grade to the gain score of each MPS student in the same grade.

For each comparison, if the MPCP student had higher growth they were given a +1; if the MPS student did better, they were given a -1; if they were tied, a score of 0 was recorded. The results are then summed across all comparisons and the result is divided by the number of comparisons. The result is Somers' d, a nonparametric measure that represents the difference between the probability that a given MPCP student will gain more than an MPS student and the probability of the opposite occurring. We also conducted the analysis on all grades pooled, since growth scores are on the same scale for each grade.[6] Table 3 reports the results of this analysis. Positive Somers'd coefficients favor MPCP students.

Table 3: Somers' d Statistics for Math and Reading Growth: 2006-07 to 2008-09

Subject/Grade	Somers' D Coefficient (s.e.)
Math 5	-.02 (.05)
Math 6	-.04 (.05)
Math 7	.09* (.05)
Math 8	-.01 (.05)
Math 10	.07 (.07)
Math All Years	.03 (.02)
Reading 5	.03 (.05)
Reading 6	.01 (.05)
Reading 7	.02 (.05)
Reading 8	.03 (.05)
Reading 10	-.05 (.07)
Reading All Years	.01 (.02)

***p<0.01, **p<0.05, *p<0.10, two-tailed. Results unweighted.

The coefficients in Table 3 should be interpreted as follows: for example, the probability that an MPCP 5th grader gained more than an MPS 5th grader in Math is .02 (or 2%) smaller than the probability of the reverse occurring. The probability that an MPCP 7th grader gained more than an MPS 7th grader in math is 0.09 (or 9%) larger than the probability of the reverse occurring. These are very small figures, and—more importantly—only one of the grade-specific coefficients (7th grade math) in Table 3 is statistically significant. That finding is perfectly consistent with the two year difference of means math results in Table 1. Still, the dominant finding is that the differences in math growth are trivial between the two sectors. As in Table 2, there are no significant differences for any grade, or overall, in reading based on Somers' d calculations.

6 See Reynolds (1977) for a further description of this procedure.

Models of Math and Reading Achievement

We are confident that the strength of our matching algorithm allows us to present the above results as valid comparisons of MPCP and MPS academic achievement growth in Year 3. However, even in the context of a random assignment study—considered by many evaluators to be the "gold standard" for internal validity—there is still analytical benefit to more elaborately modeling achievement as a function of observable student characteristics (e.g., Wolf et al., 2007, p. 33). In particular, the addition of a prior test score as a covariate can improve the precision of the estimate of a program effect. We formulate a simple model of Year 3 achievement conditioned on prior achievement, public/private school status, and student grade level:

$$\text{(eq1)} \qquad Y_{2008,i} = \beta_0 + \beta_1 C_i + \beta_2 Y_{2006,i} + \beta_3 G_i + \varepsilon_i$$

In this equation β_1 represents the impact of MPCP participation (C=1), β_2 is the impact of baseline achievement, and β_3 represents a vector of grade-specific contributions to the intercept (especially important in this context where the dependent variable by design increases by grade). With this specification, the contribution of the baseline test to the estimate of the third year test score is unconstrained in that β_2 can take any value.[7]

Although the prior achievement variable is perhaps the most important covariate, it is not the only conceivable control variable relevant to a model of student achievement. We formulate Equation 2 as:

$$\text{(eq2)} \qquad Y_{2008,i} = \beta_0 + \beta_1 C_i + \beta_2 Y_{2006,i} + \beta_3 G_i + \beta_4 X_i + \varepsilon_i$$

where β_4 represents the impact of a set of student-level characteristics, X_i, such as gender and race/ethnicity.

Previous work in different educational contexts largely confirms a negative impact of school switching on student outcomes. As elaborated below, in the context of school choice in Milwaukee, MPS students switched schools within the MPS system at a greater rate than MPCP students switched schools within the MPCP sector. Because school switching took place after our matching algorithm, we could not control for it in the original design, and its potential negative impact is important to increase the precision of our estimates. However, because switching may be "structural" in that it is induced by students reaching a terminal grade in their school, and we wish to capture voluntary switching, we include only non-structural switching in our models. We formulate Equation 3 as:

$$\text{(eq3)} \qquad Y_{2008,i} = \beta_0 + \beta_1 C_i + \beta_2 Y_{2006,i} + \beta_3 G_i + \beta_4 X_i + \beta_5 \text{Sch}_i + \varepsilon_i$$

where β_5 represents the impact of non-structural switching schools within sectors (Sch_i=1).

7 Some researchers have used differences in test scores as the dependent variable by subtracting the first year test score from the second. However, if we want to model achievement growth controlling for prior achievement, this has the effect of constraining the effect of prior achievement to equal 1.0, which empirically is not the true parameter. Thus, we favor the estimation model in Equation 1.

Finally, as noted above, some students switch sectors from the original location at baseline: either moving from MPS to MPCP or from MPCP to MPS. It is possible that these switchers differ from the vast majority of students who stay put, and that these differences are substantial enough to make comparisons of groups by their original sector misleading. To account for this possibility, we estimate one additional model:

$$\text{(eq4)} \qquad Y_{2008,i} = \beta_0 + \beta_1 C_i + \beta_2 Y_{2006,i} + \beta_3 G_i + \beta_4 X_i + \beta_5 Sch_i + \beta_6 Sect_i + \varepsilon_i$$

where β_6 is the impact of moving between sectors and β_1 is the average difference between students originally in MPCP and those originally in MPS after adjusting for whether a student switched sectors. See Table A-1 in the Appendix for descriptive sample statistics for the covariates used in Equations 1-4, and see the companion report (Cowen et al., 2010) for detailed information on the school and sector switching students.

Table 4 provides estimates of the models specified in Equations 2-4. The Model 1 column for math and reading reports results from an estimate of Equation 2. The Model 2 column corresponds to estimates of Equation 3, and so on. The results in Table 4 tell the same story as the more simple comparisons presented above. Namely, overall, there is no statistically significant average difference between MPCP and MPS students in growth in math or reading achievement. The Models 2 and 3 columns indicate that this result does not change after accounting for school and sector switching. These results are all the more supported by the estimates of the other covariates on achievement. African-American and Hispanic students score lower on average than their white counterparts—a widespread phenomenon in education research. Similarly girls do better than boys on both reading and math, but the reading advantage is greater and statistically significant. Consistent with previous research (e.g., Hanushek, Kain, & Rivkin, 2004; Lavertu & Witte, 2008) there is also a strong negative effect on math achievement associated with school switching.

Table 4: Growth Models of Math and Reading Achievement, 2006-07 to 2008-09

	Model 1 - Switching Ignored		Model 2 - School Switching Included		Model 3 - Sector Switching Included	
	Math 2008	Reading 2008	Math 2008	Reading 2008	Math 2008	Reading 2008
MPCP06	0.64	0.93	0.00	0.84	-0.52	1.09
	(1.48)	(1.64)	(1.48)	(1.64)	(1.50)	(1.66)
2006 Score	0.67***	0.75***	0.67***	0.74***	0.67***	0.74***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Nat. Am.	-4.77	7.42	-5.65	7.30	-5.87	7.40
	(6.28)	(6.41)	(6.29)	(6.41)	(6.47)	(6.33)
Asian	2.08	0.56	2.26	0.53	2.16	0.57
	(5.43)	(6.61)	(5.44)	(6.61)	(5.40)	(6.66)
Black	-10.43***	-7.58***	-10.22***	-7.56***	-10.48***	-7.43***
	(2.54)	(2.59)	(2.54)	(2.60)	(2.55)	(2.61)
Hispanic	-6.21**	-2.01	-6.43**	-2.05	-6.58**	-1.99
	(2.69)	(2.75)	(2.69)	(2.75)	(2.70)	(2.75)
Female	2.25	4.80***	2.17	4.80***	2.19	4.80***
	(1.48)	(1.68)	(1.48)	(1.68)	(1.48)	(1.68)
Switched School			-6.81***	-0.90	-6.34**	-1.12
			(2.55)	(2.90)	(2.56)	(2.91)
Switched Sector					5.46**	-2.52
					(2.75)	(3.14)
Constant	191.83***	127.06***	193.88***	127.45***	192.82***	127.92***
	(10.15)	(10.92)	(10.17)	(10.90)	(10.11)	(10.88)
N	2231	2225	2231	2225	2231	2225
R squared	0.59	0.56	0.59	0.56	0.59	0.56
F	224.02	170.99	206.25	156.81	191.63	144.57

***p<0.01, **p<0.05, *p<0.10, two-tailed. All models include indicator variables for grades 4-8, with grade 4 as the reference category; Race variables are indicator variables with "White" as the reference category. Response weights were used. Robust standard errors are in parentheses.

As discussed in the introduction, there are several ways to handle the fact that students switch sectors during a longitudinal study. In Table 4 we deal with sector switching by ignoring it, that is, by assuming that students remain in their initial sector for purposes of the analysis. That means that a student who switches from MPCP to MPS will "remain" in MPCP as measured by the MPCP indicator variable. Although this is not a randomized field trial, that assumption is standard for clinical trials in medical fields. The rationale is that in the real world people will switch medicines and conditions and it is that real world mean effect you wish to

measure. The third model in Table 4 controls for sector switching, but, because the sector-switching variable is not interacted with the MPCP variable, it does not change the fact that all achievement growth is "credited" to whichever sector each student was in at baseline, even if they are in the other sector now. Interestingly, for students as a whole that switch sectors in 2007-08, the effect is positive and significant on math. This does not change the basic conclusion from Table 4: that there are no significant differences in achievement growth in either math or reading between students in MPCP or MPS.

Although to a degree, we accept the classical assignment logic modeled in Table 4, we also acknowledge there is something different when you have a comparative observational study that is attempting to assess the relative achievement between schooling sectors. After all, MPCP switchers are receiving further instruction in MPS schools, and vice-versa. One way to test the sensitivity of our results to this problem is to compare only students who stay in the same sector for all years – in this case all three years. We have done that for the models estimated in Table 5. Descriptive statistics for covariates used in Table 5 are depicted in Table A-2.

Table 5: Non-Sector Switching (Stayer) Growth Models of Math and Reading Achievement, 2006-07 to 2008-09

	Model 1 - School switching Ignored		Model 2 - School Switching Included	
	Math 2008	Reading 2008	Math 2008	Reading 2008
MPCP06	-2.53	0.28	-3.35**	0.01
	(1.62)	(1.70)	(1.61)	(1.72)
2006 Score	0.70***	0.76***	0.70***	0.76***
	(0.02)	(0.02)	(0.02)	(0.02)
Nat. Am.	-2.83	10.70*	-3.77	10.37*
	(6.27)	(5.92)	(6.27)	(5.92)
Asian	-2.24	-2.70	-1.98	-2.62
	(5.48)	(5.70)	(5.48)	(5.72)
Black	-10.64***	-6.94**	-10.49***	-6.90**
	(2.65)	(2.71)	(2.65)	(2.71)
Hispanic	-6.63**	-1.38	-6.91**	-1.48
	(2.80)	(2.83)	(2.80)	(2.83)
Female	2.95*	4.15**	2.84*	4.13**
	(1.60)	(1.74)	(1.59)	(1.74)
Switched School			-6.85**	-2.22
			(2.77)	(3.28)
Constant	180.46***	121.36***	182.94***	122.36***
	(10.91)	(11.11)	(10.94)	(11.06)
N	1937	1927	1937	1927
R squared	0.59	0.57	0.59	0.57
F	199.31	156.69	182.91	143.70

***p<0.01, **p<0.05, *p<0.10, two-tailed. All models include indicator variables for grades 4-8, with grade 4 as the reference category; Race variables are indicator variables with "White" as the reference category. Response weights were used. Robust standard errors are in parentheses.

The results of excluding sector switchers are modest on most estimates and the covariates again conform to expectations. However, when school switching is included as a control variable, MPS "stayers" have a slightly higher achievement gain in math that is statistically significant at the .05 level. We caution readers to not over-react to that result in that this model does exclude about 300 students compared to those in Table 4, and because there are a disproportionate number of school switchers in MPS. The impact of that switching is held constant by including the switching variable in Model 2. As in all of our analyses there is no difference between MPCP and MPS students in reading.

CAVEATS

These results are limited in their explanatory power in several important ways. Nearly all concern data that are missing in some way or another, either due to study attrition or because of missing or inconsistently measured information about students who remain in the study. At this point there is little reason to believe that missing data adversely affects one or the other sectors – as shown most concretely in Appendix Table B-2 for missing students. We also will continue to backfill missing data, impute missing data on demographics, weight for missing test scores, and continue to search for missing students using a number of methods, including telephone surveys, data base searches, and even electronic tracking through the uses of "Facebook" and "Twitter."

Perhaps the most important caveat is that this study is not yet concluded with two more years remaining in data collection on achievement and perhaps longer for the attainment study. Next year's report will have the first evidence on high school completion with our 2006-07 cohort of ninth graders due to graduate (after four years) in spring of 2010. No voucher studies in the United States have reported on attainment differences between public and voucher-receiving private schools.

SUMMARY AND CONCLUSIONS

This report presents the third year analysis of academic achievement in the Milwaukee Parental Choice Program (MPCP). The analysis compares a sample of MPCP students to a sample of very similar (and in most observable ways statistically identical) MPS students. Neither a comparison of inter-sector means, nor most regression-adjusted comparisons accounting for grades, race, gender, and school-mobility indicated significant differences between the programs in terms of achievement growth in either math or reading. In general, students for whom we have two years of valid test scores appear to be performing at similar levels. Moreover, their scores appear to be similarly distributed: the gap between the highest performing students and the lowest performing students is approximately the same in the two sectors. The one exception to these findings occurs among students who have remained in the same sector for three continuous years ("stayers" in Table 5). For those students, MPS students modestly outperform MPCP students in math, but not in reading. However, the "stayers" in the MPCP and the MPS are self-selected and perhaps atypical students, and the result only becomes significant once we control for school switching, so this result should be interpreted with caution. In all other descriptive statistics and models there is no difference after two years of change in achievement growth between MPCP and MPS samples.

Throughout the report, we describe a range of cautions and caveats, with the most important being that this is only the third year of a five-year study, and that student achievement trajectories often take extended time to change. Thus, while at the present time we conclude that MPS and MPCP students appear to be approximately similar in terms of academic growth, this result may change as we incorporate the final two years of data into future analyses.

REFERENCES

Aaronson, D. (1998). Using sibling data to estimate the impact of neighborhoods of children's educational outcomes. *Journal of Human Resources*. *33(4):* 915-946.

Ballou, D. (2009). Test scaling and value added measurement. *Education Finance and Policy*. *4(4):* 351-383

Benjamini, Y., & Hochberg, Y. (1995). Controlling for the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, *57(1):* 289-300.

Cowen, J. M., Fleming, D. J., Witte, J. F., & Wolf, P. J. (2010). *School and sector switching in Milwaukee.* School Choice Demonstration Project, Fayetteville, AR: University of Arkansas.

Cullen, J. B., Jacob, B. A., & Levitt, S. D. (2005). The impact of school choice on student outcomes: An analysis of the Chicago public schools. *Journal of Public Economics*, *89(5-6):* 729-760.

Greene, J. P., Peterson, P. E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society, 31(January):* 190-213.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Disruption versus Tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics*, *88:* 1721-1746.

Heckman, J. J. (1996). Randomization as an instrumental variable." *The Review of Economics and Statistics. 78(2):* 336-341.

Howell, W. G., Peterson, P. E., Wolf, P. J., & Campbell, D. E. (2002). *The education gap: Vouchers and urban schools.* Washington, DC: Brookings Institution Press.

Lavertu, S., & Witte, J. (2008, August). *A multifacted analysis of Milwaukee charter schools.* Paper presented at the American Political Science Association Annual Meeting, Boston, MA.

Leventhal, T., & Brooks-Gunn, J. (2004). A randomized study of neighborhood effects on low-income children's educational outcomes. *Developmental Psychology, 40(4):* 488-507.

Ludwig, J., Ladd, H., & Duncan, G. J. (2001). The effects of urban poverty on educational outcomes: evidence from a randomized experiment. In *Urban poverty and educational outcomes.* Brookings-Wharton Papers on Urban Affairs. Washington, DC: Brookings Institution Press.

Reynolds, H. T. (1977). *The analysis of cross-classifications*. New York, NY: The Free Press.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70(1):* 41-55.

Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program. *The Quarterly Journal of Economics*. *113(2):* 553–602.

Rumberger, R. W. (2003). The causes and consequences of student mobility. *The Journal of Negro Education*, *72(1):* 6-21.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing 'neighborhood effects': Social processes and new directions in research. *Annual Review of Sociology*. *28:* 443-478.

Schochet, P. Z. (2007). *Guidelines for multiple testing in experimental evaluations of educational interventions, revised draft report* (MPR Reference No: 6300-080). Cambridge, MA: Mathematica Policy Research.

Witte, J. F. (2000). *The market approach to education: An analysis of America's first voucher program.* Princeton, NJ: Princeton University Press.

Witte, J. F., Wolf, P. J., Cowen, J. M., Fleming, D. J., & Lucas-McLean, J. (2008). *Milwaukee Parental Choice Program Longitudinal Educational Growth Study.* Baseline Report. School Choice Demonstration Project, Fayetteville, AR: University of Arkansas.

Witte, J. F., Wolf, P. J., Cowen, J. M., Fleming, D. J., & Lucas-McLean, J. (2009). *The MPCP Longitudinal Educational Growth Study*. School Choice Demonstration Project, Fayetteville, AR: University of Arkansas.

Wolf, P. J., Guttman, B., Puma, M., Rizzo, L. Eissa, N., & Silverberg, M. (2007). *Evaluation of the D.C. Opportunity Scholarship Program: Impacts after one year.* Institute for Education Sciences National Center for Education Evaluation and Regional Assistance: U.S. Department of Education.

APPENDIX A—Additional Tables

Table A-1: Descriptive Statistics for Variables Used in Table 4

	MPS Matched Counts	MPCP Counts
	(%)	(%)
Female	1,002	832
	(53.44)	(52.76)
White	167	144
	(8.91)	(9.03)
Black	1246*	1,013
	(66.45)	(63.51)
Hispanic	390**	399
	(20.8)	(25.02)
Asian	58*	34
	(3.09)	(2.13)
Native	12*	4
	(0.64)	(0.25)
Switched School (Non-structural)	202***	63
	(16.34)	(6.47)
Switched Sector	51***	135
	(4.13)	(13.86)

Stars indicate MPS different from MPCP statistics at ***p<0.01, **p<0.05, *p<0.10, based on a two-tailed T-Test. Demographic data on students in MPCP are provided by individual schools. There do not appear to be clusters of missing demographic data associated with particular schools. Students in grades 3, 4, 9, 11, and 12, for whom we do not have two years of achievement growth, are excluded from this table.

Table A-2: Descriptive Statistics for Variables Used in Table 5

	MPS Matched Counts	MPCP Counts
	(%)	(%)
Female	882	569
	(52.13)	(53.68)
White	159	121
	(9.40)	(11.28)
Black	1107***	608
	(65.43)	(56.66)
Hispanic	360***	316
	(21.28)	(29.45)
Asian	53	25
	(3.13)	(2.33)
Native	11	3
	(.65)	(.28)
Switched School (Non-structural)	194***	34
	(16.47)	(4.48)

Stars indicate MPS different from MPCP statistics at ***p<0.01, **p<0.05, *p<0.10, based on a two-tailed T-Test. Demographic data on students in MPCP are provided by individual schools. There do not appear to be clusters of missing demographic data associated with particular schools. Students in grades 3, 4, 9, 11, and 12, for whom we do not have two years of achievement growth, are excluded from this table.

APPENDIX B – Study Attrition

Of the original 5,424 students in the combined MPS and MPCP panels, we were unable to locate 1,202 (22 percent) in Year 3. The rate is higher for MPCP students (28 percent) compared to students who began our study in the MPS (15 percent). Some of these students may have left Milwaukee entirely, while others may have entered independent charter schools or some other educational environment outside the scope of this report. We report these figures because baseline 9th graders were a part of the original match, and will be analyzed in forthcoming studies of student attainment. However, only 3,852 students (1,926 students per sector between grades 3-8) originally matched in 2006 are eligible members of our achievement study presented in this report. Of these students, we have tracked 86 percent, with only 14 percent missing per sector after three years. That level of attrition is excellent compared to earlier studies of voucher programs (Witte, 2000; Howell, Peterson, Wolf, & Campbell, 2002).

The differences between the number of MPCP students who are missing compared to MPS is largely due to MPCP 9th graders who left MPCP schools for either MPS, other districts, other private schools, or who may have dropped out. Partly because of federal requirements to report drop out students in public schools and partly because of more consistent administrative records in MPS, students from the private sector are just more difficult to track. Fortunately we have resources to track students as part of our attainment study. That tracking is ongoing and proving quite successful in tracking down missing students, especially those in the original ninth grade. This report does not reflect those successes because we use consistent terminal dates for data collection for each year's report as fall of each year. Thus the data on attrition will change in subsequent years. Even with these numbers we are pleasantly surprised at the three-year rate of attrition for the entire study. It is much smaller than projected, and less than in other studies.

This appendix considers full sample attrition, or missing cases, including baseline 9th graders, who are not part of the achievement test study. There are two separate issues, differences in student characteristics of those who are missing from the study from those who are not; and differences in characteristics of missing students between sectors. Table B-1 addresses the first of these issues and Table B-2 addresses the second.

Table B-1: Sample Attrition Statistics 2006-07 to 2008-09

	Non-Missing Students	Missing Students
Average Mean Baseline Math	450.89**	454.69
Average Mean Baseline Reading	454.05***	459.77
%Female	51.99	51.82
%White	8.78***	13.1
%Black	67.15***	60.75
%Hispanic	20.14	21.06
%Asian	3.23*	4.05
%Native American	0.4	0.57
%Baseline Grade 3	27.86***	12.72
% Baseline Grade 4	11.21***	6.25
% Baseline Grade5	11.61***	6.87
% Baseline Grade 6	11.47***	6.17
% Baseline Grade 7	10.07***	7.49
%Baseline Grade 8	9.08	9.37
% Baseline Grade 9	18.70***	51.13

^ Stars indicate Non-missing different from missing statistics at ***p<0.01, **p<0.05, *p<0.10, based on a two-tailed T-Test.

Table B-1 indicates some racial differences between missing and non-missing students. More of the study attritors are white, while more of the students who remain in the study are black. Additionally, missing students have higher baseline reading and math scores than non-attritors. These differences are statistically significant because the sample sizes are large, however, they represent less than .1 standard deviations. Further, these differences would only be problematic for inferences about MPCP effects on achievement growth if they systematically differed between the MPCP and MPS samples. Table B-2 provides little evidence that this is the case. Among students we were not able to locate at Year 3, mean baseline reading and math scores were similar for those who began in MPS and those who began in MPCP. Missing MPS students were less likely to be female. There are some grade differences, as 9th graders made up a greater share of missing students for MPCP than MPS, and all other grades made up a smaller share of study attritors for MPCP than MPS. The current study does not include a more advanced analysis of the factors associated with sample attrition (for example, a model predicting attrition that held baseline reading and grade differences constant). We do, however, weight the observations in the outcome sample by the inverse of their probability of response, given their baseline characteristics. Incorporating such sample weights into our analysis effectively recovers in our outcome sample the careful student match that we produced at baseline (e.g., Howell et al., 2002, Appendix A).

Table B-2: MPS vs. MPCP Attrition Statistics 2006-07 to 2008-09

	MPS	MPCP
Missing Students	405 (14.85)***	793 (29.09)
Average Mean Baseline Math	453.09	452.51
Average Mean Baseline Reading	452.35	459.99
%Female	207 (51.11)***	436 (61.32)
%White	33 (8.15)	56 (7.78)
%Black	274 (67.65)	482 (66.94)
%Hispanic	79 (19.51)	161 (22.36)
%Asian	14 (3.46)	18 (2.50)
%Native American	2 (.49)	3 (.42)
%Baseline Grade 3	46 (11.36)***	34 (4.29)
% Baseline Grade 4	48 (11.85)***	32 (4.04)
% Baseline Grade5	45 (11.11)***	43 (5.42)
% Baseline Grade 6	35 (8.64)**	44 (5.55)
% Baseline Grade 7	44 (10.86)***	52 (6.56)
%Baseline Grade 8	47 (11.60)	73 (9.21)
% Baseline Grade 9	140 (34.57)***	515 (64.94)

^ Stars indicate MPS different from MPCP statistics at ***p<0.01, **p<0.05, *p<0.10

MPCP Longitudinal Educational Growth Study Baseline Report

About the Authors

John F. Witte is a Professor in the Department of Political Science and the Robert M. La Follette School of Public Affairs at the University of Wisconsin - Madison. He received his Ph.D. in political science from Yale University. He has been a fellow at the Russell Sage Foundation and at the Center for Advanced Study in the Behavioral Sciences. His latest book from Princeton University Press is *The Market Approach to Education: An Analysis of America's First Voucher Program*.

Joshua M. Cowen is assistant professor of public policy at the Martin School of Public Policy and Administration at the University of Kentucky, and is a senior analyst on a new and official evaluation of the Milwaukee voucher program He has published papers on charter school sponsorship and compliance models in school voucher randomized trials, and current research includes extensions of these models in other educational settings, including science and technology interventions in different school districts. He holds a Ph.D. in political science from the University of Wisconsin-Madison, where he was among the inaugural fellows in the interdisciplinary training program in education science, funded by the U.S. Department of Education. At the Martin School, he teaches courses in public policy and program evaluation.

David J. Fleming is an Assistant Professor of Political Science at Furman University. He received his B.A. from the University of Notre Dame in 2003 and his Ph.D. from the University of Wisconsin-Madison in 2009. He was a fellow in the interdisciplinary training program in education sciences at the University of Wisconsin. His current research project examines how education policies affect parents' political and social engagement. He teaches classes in American politics, research methods, and public policy.

Patrick J. Wolf is Professor of Education Reform and 21st Century Endowed Chair in School Choice at the University of Arkansas in Fayetteville. He also is principal investigator of the School Choice Demonstration Project. Wolf has authored, co-authored, or co-edited three books and nearly 30 articles and book chapters on school choice, special education, and public management. A 1987 *summa cum laude* graduate of the University of St. Thomas (St. Paul, MN), he received his Ph.D. in Political Science from Harvard University in 1995.

Meghan R. Condon is a doctoral candidate in the Department of Political Science, and a graduate research fellow in the Interdisciplinary Training Program in Education Sciences at the University of Wisconsin-Madison. Her dissertation is a study of students' opportunities to practice politically useful communication skills in school and the effect of those opportunities on political engagement. Her research interests include political behavior, public policy and quantitative methods.

Juanita Lucas-McLean is a Senior Study director at Westat located in Rockville, Maryland. She has more than 30 years of experience in survey research, project management, program evaluation and qualitative analysis. Ms. Lucas-McLean has managed data collection for several large-scale national surveys funded by federal agencies including the Department of Education, the Department of Labor, and the Department of Agriculture. She has a B.A. from Syracuse University and a M.S.W. from Howard University.

School Choice Demonstration Project

Department of Education Reform

University of Arkansas

201 Graduate Education Building

Fayetteville, AR 72701

479-575-6345

http://www.uark.edu/ua/der/SCDP/Milwaukee_Research.html