



Fall 2010

Gender-Separate Education: The Effects on Student Achievement & Self-Esteem on Economically Disadvantaged Public Middle School Students in Philadelphia

Heather M. O'Neill

Ursinus College, HONEILL@URSINUS.EDU

Allison Guerin

Ursinus College

Follow this and additional works at: http://digitalcommons.ursinus.edu/bus_econ_fac

 Part of the [Educational Sociology Commons](#), [Education Economics Commons](#), [Elementary and Middle and Secondary Education Administration Commons](#), [Income Distribution Commons](#), and the [Urban Education Commons](#)

Recommended Citation

O'Neill, Heather M. and Guerin, Allison, "Gender-Separate Education: The Effects on Student Achievement & Self-Esteem on Economically Disadvantaged Public Middle School Students in Philadelphia" (2010). *Business and Economics Faculty Publications*. Paper 3.

http://digitalcommons.ursinus.edu/bus_econ_fac/3

This Article is brought to you for free and open access by the Business and Economics Department at Digital Commons @ Ursinus College. It has been accepted for inclusion in Business and Economics Faculty Publications by an authorized administrator of Digital Commons @ Ursinus College. For more information, please contact aprock@ursinus.edu.

GENDER-SEPARATE EDUCATION: THE EFFECTS ON STUDENT ACHIEVEMENT & SELF-ESTEEM ON ECONOMICALLY DISADVANTAGED PUBLIC MIDDLE SCHOOL STUDENTS IN PHILADELPHIA

Heather M. O'Neill and Allison Guerin
Ursinus College

ABSTRACT

In 2003, three Philadelphia middle schools with similar demographics and failing student achievement levels were taken over by an educational management organization. Two were transformed into distinct single-sex academies within the original school buildings and a third remained coeducational. Students did not have the option where to attend, eliminating selection bias. Through funding from a *Spencer Foundation* grant, data was collected on 1,000 students for 2002-03 through 2004-05 to examine impacts of gender-segregation. We find students in single-sex schools witness greater improvements in standardized test scores, with boys gaining the most, and no differences on Rosenberg's Self-Esteem Scale.

INTRODUCTION

Numerous educational researchers, reformers, and parents believe significant changes are necessary to strengthen the current US public school system. Options, including single-sex schools, school voucher programs, charter schools, and magnet schools, are currently touted as more effective alternatives to the traditional coeducational public school system. The No Child Left Behind Act (NCLB) of 2001 stipulates poor performing schools within chronically low achieving school districts may be subject to state takeovers or management by private organizations, which in turn may choose an alternative to the traditional coeducational public school model.

In 2002, following years of disappointing academic performance and armed with provisions in the NCLB Act, Pennsylvania government officials created the School Reform Commission (SRC) to run the Philadelphia school system. The SRC developed the 'diverse provider' model wherein it hired seven private managers and two universities to run the 45 poorest performing primary and middle schools. Another 37 poor-performing schools retained Philadelphia district management albeit with additional funding, including 21 schools granted additional staff. The SRC continues to oversee the 180,000 students in the 270 public schools in grades K-12 in Philadelphia.

One of the chosen private education management organizations (EMOs) took charge of two middle schools in 2002-03. The first school was immediately segregated by gender into two academies within the same building. By 2005-06, that school was transformed into a boys-only public academy. The second middle school was initially slated to be segregated by gender, but setbacks precluded it, thus it remained coeducational. Not until 2005-06 did that school become a girls-only public academy. In 2003-04, the same EMO was given a third contract, transforming another middle school into two single-sex academies within the same building. The same

curriculum was taught in all three schools. Students remained in the middle school in which they previously matriculated, thus did not have a choice as to which school to attend. This provided a natural experiment devoid of self selection to discern changes in academic performance and self esteem attributable to differences in single-sex versus coeducational school or classroom settings. Other studies on single-sex education generally focus on tuition-charging schools wherein parents and students choose the environment, thus self selection bias occurs, and parental income rather than school setting generally impacts academic achievement. Thus, eliminating selection bias, having similar curriculums taught in all three schools, and focusing on failing schools all populated with economically poor students creates propitious conditions for this study. The focus of this paper is whether single-sex education in economically-disadvantaged, public middle school settings leads to higher levels of academic achievement and student self-esteem.

GENDER-SEPARATE EDUCATION IN THE US: THEORY AND EVIDENCE

Although the first public schools in the US were exclusively for males, the eventual educational opportunities for females led to a national standard of coeducational public schools by the middle of the nineteenth century (Tyack & Hansot, 1990). Coeducational schools were seen as more economically efficient, providing a major impetus to their predominance (Riordan, 2002). Riordan contends this historical preference to coeducation creates a ‘protective halo’ for the institution, allowing it to exist without proving its efficacy in student outcomes, while placing a burden of proof on single-sex institutions “to show greater effectiveness” (2002, p.11).

When compared to other countries, only a small number of public single-sex programs have been implemented for trial periods in the United States. Single-sex schooling experiments designed for the US were generally abandoned because of fears the programs would be challenged in court under Title IX requirements of equal funding for males and females. The US Department of Education’s recent modifications of Title IX enable more public school districts to offer gender-separate education. According to the National Association for Single Sex Public Education, 51 single-sex public schools existed in 2006 compared to only three in 1995 (Philadelphia Inquirer, 2006).

Differences regarding the most efficacious method of education are numerous and ongoing. Supporters of coeducation, similar to those who supported racial desegregation, believe that by placing all students in the same environment, an equality of opportunity is created for all students. Sexism and gender discrimination, whether in a school, work, or social setting, can be mitigated through educational integration of the sexes. All students learning together in the same classroom are presented with the same educational experiences and resources, thus promoting equality in outcomes via equality in inputs.

Much rationale supporting single-sex programs is based on the belief that coeducational schooling environments create distractions for boys and girls, impeding their ability to learn at their highest ability. Through student interviews, Streitmatter found girls perceived boys to be ‘academic barriers’ to their learning, whereas boys saw girls as “sexual distractions” in the classroom (2002, 224). According to the National Coalition of Girls’ Schools, single-sex schools are better suited for young females because they create an environment that supports risk-taking, leadership, academic excellence, personal development, and better preparation for college (Ransome & Moulton, 2001). As reported by researchers Monaco and Gaier, girls in single-sex schools have higher self-esteem and are encouraged to achieve beyond the stereotypical roles associated with women (Watson et al, 2002). Similarly, Gilson (2002) argues females, being more socially oriented than boys, should profit in single-sex classes due to less social noise in the classroom.

Quantitative studies focusing on females show mixed results. Early studies using Catholic school student data show females in single-sex schools gain more confidence, build greater self-

esteem and have higher test scores than girls in coeducational schools (Lee & Bryk, 1986; Lee & Marks, 1990; Riordan, 1985). Later studies using Catholic school data yield contrary results. For instance, LePore & Warren (1997) found no significant difference in achievement test scores between single-sex versus coeducational schools. Similarly, using data for students attending tuition-charging, independent schools, Gilson (1998) found no statistically significant differences in math attitudes or mathematical achievement for girls in the single-sex versus coeducational schools, contrary to her expectations.

Riordan (2002) contends income level is the defining variable in determining the efficacy of single-sex education; lower income students gain more from gender-separate education, regardless of gender or race. He argues, “the research is ‘exceedingly persuasive’ in demonstrating that single-sex schools are effective in terms of providing both greater equality and achievement, especially for low-income and working class students, most particularly for African-American and Hispanic-American boy *and* girls.” (2002,13). Single-gender schools impose a more pro-academic environment and seriousness of purpose that benefits under-privileged students who may lack this discipline from home. He claims 70% of the difference in test scores between single-sex and coed schools for black and Hispanic students can be attributed to differences in adolescent subculture, coursework, curriculum, and the amount of homework. Thus, Gilson’s (1998) insignificant findings are not surprising since few economically disadvantaged youths attend tuition-charging schools. Moreover, the increased affluence of girls attending Catholic schools beginning in the 1990’s accounts for the mitigation of positive impacts for single-sex education seen in the 1970’s and 1980’s (Riordan, 2002).

Support of single-sex schools targeting African American boys comes from the belief that a constant flow of negative stereotypes concerning African American men, coupled with the low expectations teachers exhibit towards African American students, are present in most of the coeducational schools in America. Single-sex schools are effective in creating an environment in which these boys have strong positive role models who encourage a higher level of self-esteem and provide students with concrete examples of successful African American males in society (Singh et al, 1998 and Riordan, 1994). These impacts, in turn, lead to greater academic outcomes. Singh et al (1998) studied 90 urban fifth grade boys and girls in both classroom settings and found boys in the single-sex classes had higher grades but lower achievement test scores, whereas girls in the single-sex classes outperformed their coed counterparts in six of eight achievement tests and class grades. Single-sex school attendance rates, however, were significantly higher for boys and girls.

Overall, the findings on the benefits of single-sex schools on student achievement and self-esteem are mixed. The inconclusiveness of these results suggests additional research is needed, especially in the public school sector, since these programs are so rare. Additionally, it is worthwhile concentrating on middle school students, those for whom sexual differences become more apparent as they enter puberty, since quantitative studies for these students are lacking. The natural experiment conducted in Philadelphia provides these research opportunities.

THE PHILADELPHIA EXPERIMENT

Philadelphia’s SRC placed three coeducational, underperforming middle schools under the control of a leading for-profit EMO beginning in 2002. By fall 2003, two schools segregated students by gender by creating two academies within the school building, while the third school remained coeducational. The EMO used the same curriculum in all three schools with a special emphasis on hands-on science education; there were no stated differences in curriculum based on the school setting or any discussion of specifically addressing self-esteem issues. Additionally, all students were in classes of similar size regardless of the size of the school. As of this writing, all three are still managed by the EMO. For the purposes of this study, it is advantageous the third

school remained coeducational and used the same curriculum, since that school can serve as the control.

The three schools service predominantly African American students from economically disadvantaged families. Table 1 highlights socioeconomic and academic achievement at the three schools for 8th grade middle school children for 2002-03 (Philadelphia Inquirer, 2003). Generally, across the schools, more than 60% of students performed below the basic level of proficiency in reading and more than 70% below proficiency in math. The percent of African American students in each school varied from 87% to essentially 100%, and these percentages are roughly the same today. The percent of students eligible for free school lunches, a measure of economic disadvantage, varied between 73% and 86%. By 2008-09, all three had about 84% eligible, compared to the city average of 76% (School District of Philadelphia, 2009). This similarity in socioeconomic backgrounds helps focus differences in achievement effects to gender separation, not economic background.

Table 1: Academic and Demographic Characteristics of 8th Grade Students in 2002-03

	Single-Sex #1	Single-sex #2	Coed
	Two academies in one building in 2002-03.	Two academies in one building in 2003-04.	Remained coed in 2003-04 and 2004-05.
	Became all-boys school in 2005-06 for grades 6-12.	Remains as such today for grades 5-8.	Became all-girls school in 2005-06 for grades 7-12.
	Middle school 2003-04 enrollment=368.	Middle school 2003-04 enrollment=582	Middle school 2003-04 enrollment=517
% students eligible for free/reduced lunch program	86	73	82
% students who are African American	99.7	86.7	98.4
MATH*			
% Advanced	0	1	0
% Proficient	3	4	3
% Basic	10	22	16
% Below Basic	87	73	81
READING*			
% Advanced	0	1	1
% Proficient	8	14	9
% Basic	23	25	22
% Below Basic	69	61	69

*Scores from the 2002 Pennsylvania System of School Assessment Exam for 8th Grade Students (Pennsylvania School Test Scores, 9-10)

Since the implementation of these reform efforts, the RAND Corporation and Research for Action published an official report on the status of academic achievement in Philadelphia's public schools (Gill et al, 2007). According to RAND, from the 2001-02 to 2005-06 school years, there has been a 20% increase in students reaching proficiency in eighth-grade reading, and a 19% increase in students reaching proficiency in eighth-grade math. However, these results do not take into account statewide gains over this same period of time. When Philadelphia schools ranked in the lowest quartile in achievement scores were compared to the other schools in the state in the lowest quartile, it was found Philadelphia's schools outgained the others in eighth-grade reading scores by a statistically significant margin, but there were no statistically significant differences for eighth-grade math scores. Comparing all the EMO schools (not just the three in this study) to all other public schools in Philadelphia, there were no statistically significant differential effects on student achievement scores in reading or math after four years. The report acknowledges, however, the fact that EMOs were given the lowest-achieving schools in the district, then at the very least, even if their gains did not exceed those of the district at large, they did improve at a consistent pace with the rest of the schools.

Data for this study come from several sources. The SRC provided end-of-the-year achievement test scores for all students in the three middle schools for the three academic years 2002-03 through 2004-05. The total test score, called the student standardized TerraNova total test score, evaluates the combination of reading, math, science, and language abilities and are used to monitor academic achievement. The test score "characterize[s] proficiency in absolute terms without making comparisons to the proficiency or growth of students in a reference group. Higher scale scores indicate higher proficiency." (CTB/McGraw-Hill, 2006,1). Thus, the scores allow for direct comparisons of students and schools. The SRC also provided gender and school attendance records for all students in the three schools.

To gather information on student self-esteem, survey data derived from student interviews from the three schools were made available through a Spencer Foundation grant. Toward the end of the 2003-04 academic year, random students selected from all three schools responded to the survey: 291 from the coeducational school and 211 from the single-sex schools. A second survey, administered one year later in spring 2005, included 166 different students from the coeducational school and a different 336 from the two single-sex schools combined. The data are not longitudinal; the same students were not tracked via the survey over time. The survey questions span several areas. Some address the student's home life and parental work experiences. Others ask about activities and time spent on them. Student post-high school educational and occupational aspirations and impressions about self image are also solicited.

A widely used self-esteem index created by Rosenberg allows a quantitative measure of a "positive or negative orientation to oneself; an overall evaluation of one's worth or value" using the survey questions (Rosenberg, 1989, 1). The Rosenberg Self-Esteem Scale is based on ten questions answered on a four point scale (zero to three) ranging from "strongly disagree" to "strongly agree." The scale scores range from 0 to 30, with 30 indicating the highest self-esteem score possible. Reverse scored questions were renumbered when compiling the self-esteem score. Table 2 indicates the questions included in the Rosenberg Self-Esteem Scale, with the reverse scored questions indicated by an asterisk.

MODELS

Academic Achievement Model

Based on the previous literature listed above, we posit middle school student academic achievement is determined by demographic traits and socioeconomic family background, student behaviors, and school setting – either coeducational or single-sex. Equation (1) represents a

Table 2: Rosenberg’s Self-Esteem Scale Questions

		Strongly Agree	Agree	Disagree	Strongly Disagree
1.	On the whole, I am satisfied with myself.	3	2	1	0
2.*	At times I think I am no good at all.	0	1	2	3
3.	I feel that I have a number of good qualities.	3	2	1	0
4.	I am able to do things as well as most other people.	3	2	1	0
5.*	I feel I do not have much to be proud of.	0	1	2	3
6.*	I certainly feel useless at times.	0	1	2	3
7.	I feel that I’m a person of worth, at least on an equal plane with others.	3	2	1	0
8.*	I wish I could have more respect for myself.	0	1	2	3
9.*	All in all, I am inclined to feel that I am a failure.	0	1	2	3
10.	I take a positive attitude toward myself.	3	2	1	0

multiple regression equation for academic achievement, ACHIEVEMENT, with vectors for the sets of regressors. A student’s total TerraNova score is used to measure achievement, rather than focusing on specific area of learning, since the goal is overall improvement in a student’s academic success. The vector for family background and demographics, DEMOGRAPHICS, serve chiefly as control variables, since a student’s family background affects achievement by enabling and promoting academic effort. Two post-graduation education aspiration variables impacting achievement are included therein: the education needed for a student’s likely occupation and whether a student plans to attend college. Students intent on having an occupation requiring a college degree or planning to attend college are expected to achieve higher test scores, since higher test scores increase the likelihood of college admittance. It is possible, however, that higher achievement leads to the aspiration to attend college, thereby creating an endogenous relationship that biases the results. One included parental variable, whether a student lives with two parents, serves as a control variable. The other, whether a student’s grades are consistent with parental expectations is expected to increase achievement. The presumption of good grades demanded by parents followed by students getting them suggests greater effort and achievement. Parental employment status and educational level are used in other studies (Thompson, 2003; Ehrenberg et al, 1995; Jimenez and Lockheed, 1988), but too many missing observations preclude their use here. Since all three schools have significant African-American populations, race is not included as an independent demographic variable, but gender is.

$$ACHIEVEMENT_{it} = \beta_0 + \beta_d * DEMOGRAPHICS_{it} + \beta_b * BEHAVIOR_{it} + \beta_a * ACHIEVEMENT_{it-1} + \beta_s * SELF-ESTEEM_{it} + \beta_{ss} * SINGLESEX_{it} + \beta_f * SINGLESEX * SEX_{it} + \epsilon_{it} \quad (1)$$

where t = academic year, i = student, and ϵ_{it} represents the stochastic error.

BEHAVIOR captures school behaviors consistent with trying to improve academic achievement, such as fewer school absences and more hours spent on homework per day. Higher SELF-ESTEEM, measured by Rosenberg’s Self-Esteem Scale, is also expected to lead to higher achievement, i.e., $\beta_s > 0$.

Using previous year's test scores as a regressor serve two purposes. First, they allow for the value added in achievement during the current year by holding previous achievement constant. Second, omitted variable bias associated with unobserved traits in E_{it} , such as ability, correlated with SELF-ESTEEM may be mitigated. Growth in unobserved traits can still create bias, but unchanging levels in them do not affect the value-added, as discussed by Jimenez and Lockheed (1988).

It is hypothesized that students in the single-sex middle schools will perform better academically than those in a coeducational setting, $\beta_{ss} > 0$, but the differential impact for males and females is zero, $\beta_1 = 0$. Since the schools serve students from economically disadvantaged areas, these premises consistent with Riordan's (2002) hypotheses are readily testable. Aforementioned studies, such as Lee & Bryk (1986) and Lee & Marks (1990) suggest females in single-sex schools outperform males ($\beta_1 > 0$), suggesting the "less distraction environment" impacts girls more than boys. However, when these NCLB schools were designated as failing to meet certain standards, the EMO was partnered with them for the purpose of increasing the achievement level of all these schools' students, not necessarily to enhance one gender's achievement more than the other. The EMO's initial intent was clearly to create single-sex learning environments, which they eventually did in all three schools, thus the EMO must have believed $\beta_{ss} > 0$. If the sign on the interaction term, β_1 , is not zero, it is likely an unintended consequence suggesting something about the power of single-sex education on the different genders.

Self-Esteem Model

Equation (2) indicates the multiple regression model for self-esteem with Rosenberg's Self-Esteem Scale as the dependent variable. The demographic and family background variables serving as control variables in (2) are not the same as in (1), except for gender. Two parental variables are included: whether the student behaves as parents expect and whether parents share similar futuristic goals as their student. Positive answers to both are expected to increase student self-esteem, since parental expectations are aligned with the student, thus creating a positive outlook with parental affirmation of this sense of self. Another trait, having high occupational aspirations, which is measured by whether a college degree is needed for the student's dream job, is likely to increase self-esteem. Students with big dreams are thought to have more self-esteem. The last demographic variable is church membership, serving as a control variable.

$$\text{SELF-ESTEEM}_{it} = \gamma_0 + \gamma_d * \text{DEMOGRAPHICS}_{it} + \gamma_a * \text{ACHIEVEMENT}_{it} + \gamma_{ss} * \text{SINGLE-SEX}_{it} + \gamma_1 * \text{SINGLESEX} * \text{SEX}_{it} + E_{it} \quad (2)$$

where t = academic year, i = student, and E_{it} represents the stochastic error.

Students attending single-sex middle schools are hypothesized to exhibit a higher level of self-esteem than those attending coeducational middle schools, $\gamma_s > 0$, and boys gaining greater self-esteem than girls in such settings, $\gamma_1 < 0$. Ceteris paribus, boys and girls develop greater levels of self-esteem when separated from the opposite gender due to fewer distractions and more opportunities to gain confidence without the pressure of embarrassment from the opposite gender. The literature focused on African American boys suggests boys benefit more. One should note that previous self-esteem scores are not used as regressors, as in the achievement model (1), due to not having two self-esteem indices on any student. Their exclusion may bias the results.

Academic achievement and self-esteem may be simultaneously determined, as suggested by the literature; higher self-esteem levels lead to greater test scores and higher test scores enhance self-esteem. If so, estimating the structural models (1) and (2) via ordinary least squares given concurrent measures of the dependent variables leads to biased results, thus necessitating the use

of two stage least squares (2SLS) with reliable instruments. In theory, both structural models can be estimated, since the rank conditions hold due to different exogenous variables in (1) and (2); several of the instruments appear in both structural equations, while some do not, allowing for proper identification of each structural equation. If the variables are not simultaneously determined, ordinary least squares or instrumental variable estimation techniques are relevant.

DATA

Table 3 presents the descriptive statistics for the data, broken down by gender and school setting for those students interviewed in spring 2004. Despite over 500 surveys taken, incomplete survey responses and missing test scores led to a sample size of 310 for all variables, except Test Scores in 04-05. Numerous students, more so boys than girls, with test scores '03-'04 did not have test scores recorded for '04-'05, leading to the smaller sample sizes in Test Scores 04-05. Though not shown, the overall average '03-'04 test score was 655.3 with a standard deviation of 31.9, while the self-esteem index averaged 22.57 with a standard deviation of 4.69.

Table 3: Mean Values for Variables for Spring '04 Respondents (n is sample size)

Variable	Single-Sex		Coeducational	
	Male (n=54)	Female (n=71)	Male (n=61)	Female (n=124)
Test 02-03	646.7 [33.3]	656.5 [27.0]	646.4 [27.6]	644.8 [25.5]
Test 03-04	660.8 [37.4]	663.5 [27.9]	650.2 [33.7]	650.7 [29.4]
Test 04-05	675.4 ^a [38.4]	677.2 ^b [27.1]	661.6 ^c [33.4]	660.4 ^d [29.7]
Self-Est 03-04	23.46 [4.4]	22.32 [4.8]	22.5 [4.5]	22.37 [4.9]
Church 03-04	0.52	0.62	0.54	0.58
TwoPar 03-04	0.46	0.35	0.34	0.29
EdThink 03-04	0.24	0.56	0.20	0.50
EdHope 03-04	0.31	0.62	0.27	0.60
GoColl03-04	0.037	0.099	0.114	0.064
PAlign 03-04	0.69	0.66	0.70	0.69
PGrade 03-04	0.63	0.69	0.56	0.69
PBehave 03-04	0.74	0.76	0.56	0.69
Absences 03-04	15.9 [10.6]	13.22 [10.1]	16.7 [12.4]	14.69 [11.6]
HomeWk 03-04	0.35	0.19	.031	0.23

a (n=42) b (n=68) c (n=48) d (n=110) [standard deviations in brackets]

Perusing Table 3 indicates improvements in average Test Scores from '02-'03 to 04-'05 of 29 and 21 points for males and females, respectively, in single-sex environments. The improvements in the coed settings were 15 for males and 16 for females. Yet, if one looks at levels without an eye to value-added, one sees females in single-sex schools have the highest average score (677.2) in '04-'05 but only a two point edge over boys in these schools (675.4). Both genders have lower test score levels in the coed schools with little difference across genders. These observations reinforce the need for a value-added model of achievement.

Church membership, Church, ranged between 52% and 62% overall. A greater percentage of boys lived with two parents, TwoPar, than girls living with two parents in both school settings. Fifty percent or more of girls indicated their likely occupation requiring a college degree, EdThink, whereas only 20% of boys did. For the education needed for the occupation they would like to have, EdHope, as opposed to the one they would likely have, the percentages rose modestly for both males and females. Only 3.7% of males in single-sex schools planned to attend

either a two year or four year college, GoColl, compared to 11.4% of boys in the coed schools. For girls, however, the trend was reversed. The percent of the students who believed they shared similar future goals as their parents, PAlign, was about 70% across the board. For grade expectations, PGrade, only 56% of boys in coed schools believed their grades were consistent with their parent's expectations compared to 63% for boys in single-sex schools. A greater percentage of students in single-sex schools, above 70%, believed they met their parent's expectations for behavior, PBehave, than children in coed schools. Average school absences in '03-'04, Absences, were slightly lower for both genders in the single-sex schools, roughly fourteen days, compared to fifteen days in the coed schools. Weekly hours spent on school work were coded into percentiles with HomeWk is defined being in the top 25th percentile in hours spent on school work per week, which translated into 4 or more hours per week. Thirty five percent of the boys in the single-sex schools were in the top quartile compared to only 19% of the girls; male predominance was less in the coed schools.

Table 4 is an abbreviated table for those students receiving the survey in '04-'05. The sample sizes are noticeably smaller and the average test scores for each year are markedly lower than those for the '03-'04 respondents, although it is not clear why this is the case. Since the respondents in both years tended to be eighth graders, it is not due to surveying different grades of middle-schoolers. Fortunately, the value added approach in (1) adapts for the differences in score levels when using pooled data. The improvements in average test scores by gender between '03-'04 and '04-'05 are greater for students in the coed settings, contrary to the results for the previous year's respondents, albeit with lower sample sizes.

Table 4: Mean Values for Spring '05 Respondents (n = sample size) [std. dev.]

Variable	Single-Sex		Coeducational	
	Male (n=30)	Female (n=51)	Male (n=32)	Female (n=30)
Test 03-04	642.1 [31.5]	656.1 [27.4]	632.0 [31.3]	637.8 [35.9]
Test 04-05	651.3 [38.1]	667.1 [28.0]	644.2 [35.4]	659.7 [27.6]
Self-Est 04-05	21.90 [4.6]	23.31 [4.3]	23.13 [6.5]	22.26 [4.8]

EMPIRICAL RESULTS

Achievement Scores

First, we concentrate on the '03-'04 survey respondents because this is the first year that two schools had single-sex academies within their buildings and the sample size is larger than for the '04-'05 respondents. Estimating (1) and (2) by 2SLS for the '03-'04 respondents, which is appropriate if the endogenous variables are simultaneously determined and robust instruments are available for identification, yields the results in column two in Table 5 for the achievement model (1) and column two Table 6 for the self-esteem model (2). The insignificant coefficient on self-esteem in Table 5 makes one question not only its significance but its inclusion as being endogenously determined with the test score. Using the Hausman technique to test whether self-esteem is endogenous in (1), as outlined by Wooldridge (2009, 528), yields a p-value of 0.92 on the predicted residuals coefficient. A statistically significant p-value, say one less than .10, would enable us to conclude self-esteem is endogenous. Based on the high p-value of .92, we have insufficient evidence of whether self-esteem is endogenous, which is espoused by theory.

Table 5: Estimations of Achievement, Test 03-04

Variable	2SLS Par. Est. (p-value)	IV Par. Est. (p-value)	OLS Par. Est. (p-value)	Pooled OLS ^a (p-value)
Intercept	89.90 (0.0016)	90.18 (0.0015)	90.66 (0.0009)	97.69 (0.0001)
Test 02-03	0.85 (0.0001)	0.85 (0.0001)	0.85 (0.0001)	0.84 (0.0001)
Self-Est 03-04	0.15 (0.9235)	0.21 (0.8965)	0.30 (0.2148)	0.30 (0.1659)
Single-Sex	9.00 (0.0193)	8.96 (0.0199)	8.89 (0.0144)	4.77 (0.1274)
Female	-1.18 (0.7055)	-1.18 (0.7072)	-1.16 (0.7906)	1.85 (0.5246)
SingleSex*Female	-6.77 (0.1800)	-6.70 (0.1846)	-6.59 (0.1569)	-3.18 (0.4215)
EdThink	7.63 (0.0023)	7.65 (0.0023)	7.69 (0.0016)	5.26 (0.0160)
PGrade	6.37 (0.0819)	4.72 (0.2320)	6.11 (0.0113)	4.58 (0.0285)
GoColl	-0.88 (0.8436)	-0.0894 (0.9843)	-0.75 (.8591)	-0.33 (0.9262)
TwoPar	2.85 (0.2270)	2.85 (0.2266)	2.84 (0.2270)	3.13 (0.1345)
HomeWk	-0.27 (0.9134)	-0/35 (0.8911)	-0.28 (0.9094)	2.33 (0.3069)
Absences	0.011 (0.9272)	0.008 (0.9494)	0.005 (0.9574)	0.09 (0.2678)
Adj-R ²	0.639	0.6402	0.6404	0.5996
Sample Size	310	310	310	422
Dependent Mean	655.3	655.3	655.3	656.3

^a The dependent variable is test 03-04 or test 04-05, depending on the response year. Similarly, the previous test scores are test 02-03 and test 03-04.

If theoretically self esteem does not interact endogenously with achievement, meaning self-esteem is exogenous in (1), and the Hausman test fails to indicate endogeneity as shown above, then 2SLS is less efficient than OLS and OLS is preferred. However, before proceeding to OLS estimation, a case can still be made for omitted variable bias, which warrants instrumental variable (IV) estimation to eliminate the bias. From the value-added model, if changes in unobserved traits occur and they are correlated with self-esteem, IV estimation is recommended. Whether such growth occurs is not certain. If growth in motivation or ability occurs and it's positively correlated with self-esteem, then the coefficient on self-esteem is positively biased using OLS. The criterion for inclusion as an instrumental variable is that the variable be correlated with self-esteem but not with unobserved growth factors in the error term. Three chosen instruments, all found to be statistically significant predictors of self-esteem (joint significance p-value=.0073 and individual p-values less than .10), yet not necessarily correlated with growth in ability or motivation are PAlign, PBehave and EdHope. For example, it is not clear that an alignment of parental expectations vis-à-vis a student's future is related to the student's growth in ability or motivation, yet it does ably predict self-esteem. The IV estimation results are shown in the third column of Table 5. Again, a Hausman test can discern if IV is preferred over OLS if the

Hausman test supports endogeneity. If self-esteem is not endogenous with the error term, OLS provides better results. This second Hausman test yields a p-value on the predicted residuals of 0.95, again suggesting a lack of significance for endogeneity. We report the OLS estimates, which do not exhibit heteroskedasticity, in column four.

The estimates from the three estimation techniques are fairly similar in both levels of significance and size of impacts, as shown in columns two through four. All p-values are for two-tailed hypothesis tests regarding no impact of each predictor individually on achievement. Halving the p-values shows one-tailed levels of significance and any such value at or below .10 is noted as indicating statistical significance. Several predictors are statistically significant and they have the expected sign. Without loss of generality, the OLS estimates are presented. The previous test score coefficient of .85 suggests a one-point test score increase in '02-03' leads to .85 point higher score in '03-'04, *ceteris paribus*. Students with college degrees needed for the occupation that they think they will most likely obtain are expected to have higher test scores of 7.69 points, which represents one fourth of a standard deviation from the mean. Consistency in parental and student grade expectations increase scores by 6.11 points on average. The Single-Sex school coefficient indicates higher test scores of nearly 9 points for students in single-sex environments. The interaction term SingleSex*Female coupled with Single-Sex yields a joint level of significance of .03. Thus, boys in single-sex schools are expected to outperform their male counterparts in coed schools by 9 points, whereas girls in single-sex are only expected to outperform their counterparts by 2.30 points, holding previous test scores constant. These improvements are attributable to the value-added by the school setting, *ceteris paribus*. The insignificant p-values for Female alone and with the interaction term imply males and females do not show differential increases in test scores in the single-sex or coed school settings. Therefore, single-sex education increases test scores for both genders relative to their peers in coed schools, but not differentially for boys or girls in single-sex schools, as hypothesized.

Contrary to expectations, self-esteem is not statistically significant. Interestingly, if one omits previous Test 02-03, self-esteem is statistically significant with a coefficient of 1.45, but all the school setting and gender variables are insignificant. This is not surprising given the expected positive bias due to omitting Test 02-03. Failing to properly include Test 02-03 would lead to highly improper conclusions regarding the efficacy of the single-sex school option.

The last column in Table 5 presents pooled cross sectional results using both years of data. Although pooling allows for a greater sample size, it muddies the analysis because it assumes the changes in achievement are the same each year, even though the second year's data coincides with the second year of the gender-separate experiment. Unfortunately, we cannot discern changes for a given student over time to see if the achievement changes are permanent, because the data are not longitudinal. One Chow test indicated both years had similar structural models for (1) and a second Chow test suggested the intercept alone was not different across the years. Joint significance tests were performed for the interaction terms with gender and school type and neither were significant. Additionally, a Hausman test once again showed a lack of significance for the endogeneity of self-esteem. The results suggest males in single-sex schools outperformed their peers at coed schools by 4.77 points ($p\text{-value} = 0.1274$) and no difference for females across schools. The decline in significance and magnitude for β_{ss} suggest boys in single-sex schools outperform their coed peers, albeit by a lower level and with less confidence in that level than indicated by OLS in column four.

Self-Esteem

Once again, we concentrate on the first year in which the two schools were single-sex, '03-'04, and present the results for predictors of Self-Est 03-04 in Table 6. The 2SLS results in column two indicate a ten point increase in the concurrent test score increases the self-esteem

index by .49, about a 2% increase. The Hausman test yields a p-value on the predicted residual term of 0.05, thus the Self-Est 03-04 and Test 03-04 are endogenously determined in (2). An increase in achievement increases self-esteem slightly, but higher self-esteem does not statistically significantly improve achievement. Students who believe they behave as their parents expect show a 1.09 point increase in self-esteem, ceteris paribus. Students who share similar expectations for their future as do their parents have an expected higher self-esteem index of .87 points. If a student's dream job requires a college education, they are expected to have a lower self-esteem index of 1.34 points. Contrary to expectations, school type and gender have no impact on self-esteem. Although single-sex schooling can lead to higher test scores, shown in Table 5, which in turn can raise self-esteem, shown in Table 6, enhancing self-esteem does not seem to be occurring in isolation in gender-separate classes any differently than in coed classes. Likewise, there is no apparent difference in self-esteem levels across gender. Without longitudinal data we cannot comment on how self-esteem is changing over time per student.

Table 6: Estimations of Self-Esteem, Self-Est 03-04

	2SLS Par. Est. (p-value)	Pooled OLS ^b (p-value)
Intercept	-10.26 (0.1324)	-5.05 (0.3765)
Test 03-04	0.049 (0.0001)	0.043 (0.0001)
Single-Sex	0.31 (0.7140)	-0.19 (0.7917)
Female	0.14 (0.8494)	-0.55 (0.4080)
SingleSex*Female	-1.04 (0.3387)	-0.15 (0.8687)
Church	0.20 (0.7034)	0.16 (0.7255)
PAlign	0.88 (0.1183)	1.03 (0.0365)
EdHope	-1.34 (0.0148)	-0.68 (0.1668)
PBehave	1.09 (0.0581)	0.63 (0.1898)
Adj-R ²	0.0847	0.0498
Sample Size	310	422
Dependent Mean	22.57	22.76

^b The dependent variable is test 03-04 or test 04-05, depending on the response year. Similarly, the previous test scores are test 02-03 and test 03-04.

Pooled data Chow tests revealed no structural model or intercept differences for the two years' models. The 2SLS results for the pooled sample appear in the third column. Again, the Hausman test found test score and self-esteem are endogenously determined in (2) with a p-value of .08 on the predicted residuals. The results are weaker than for the 2SLS '03-'04 model as only PAlign and Test Score are statistically significant. The impact on self-esteem from Test Score is similar at .043 and there is no difference across males and females.

CONCLUSION

The applicability of this study to the national school reform debate is notable, especially as alternatives to the traditional coeducational public school setting become more popular. The natural experiment that took place in Philadelphia provides evidence that economically disadvantaged middle school boys show greater increases in academic achievement test scores than boys attending coeducational schools, *ceteris paribus*. The performance differential for girls is not as apparent, and when it occurs it is a smaller differential. Since the curriculum was the same in all the schools and self-selection was not an option, the results have great appeal.

Self-esteem formation does not appear to be taking place any differently in single-sex versus coeducational settings. Without longitudinal data we cannot discern the value-added in self-esteem, but our results do find no apparent differences in levels in self-esteem across schools or gender. Self-esteem does not impact test scores, contrary to the original hypothesis, but test scores do affect self-esteem modestly. Since the EMO running the schools did not reveal any programmatic policies to address self-esteem, we have no reason to believe the EMO attempted to build self-esteem differently in either school setting, just as there was no overt policy to improve test scores differentially. Therefore, while higher test scores improve self-esteem, the higher test scores are being generated by reasons other than higher self-esteem.

We also learn that links between the student and parents in terms of similar expectations for grades, behavior and future goals impact achievement and self-esteem. Alignment of grade expectations improves test scores, while the other parental nudges increase self-esteem. Not surprisingly, stronger parental support and direction have the expected effects, and developing even stronger links may be feasible through parent/teacher/student interactions.

The three educational aspiration variables show mixed results, in part due to deficiencies in the data. The relatively small percentage of students who state they want to attend either a 2-or-4-year college someday may be the root of the insignificant coefficient, namely too little variation in the regressor. On the other hand, the types of likely or dream occupations listed by students and whether a college degree is required for them, is skewed toward too little requiring college for boys. Numerous boys list professional athlete as either their likely or dream job, which were coded as not requiring a college degree. Yet these boys exhibit higher levels of self-esteem, most likely due to their perceived athletic prowess, and not necessarily a devotion to doing well on achievement tests. When asked about their most likely occupation, those requiring a college degree do exhibit significantly higher test scores. A dose of reality and greater awareness of occupations and the education needed for them would serve these students well in promoting achievement.

Two behavioral variables, absences and hours spent on homework, which are subject to more control by teachers and administrators, do not impact test scores. This unexpected result for homework is probably due to the relatively few hours of homework undertaken per week. About 75% of the students reported less than four hours per week. It is not clear if greater hours are demanded and either ignored by students or not reinforced by parents or if relatively few hours are required, but the low figure could provide the disconnect with homework effort and test scores. It is not clear why being absent from school also does not deleteriously affect test scores.

Two chief drawbacks of the study are the data are not longitudinal and there are only two years of self-esteem data and three years of test scores available. Tracking changes over a longer time period, even if different students were interviewed each year, would be propitious to see if the initial increases in test scores from single-sex academies continue over time. Similarly, following a cohort of students over time would show whether any initial test score gains were fleeting or just the beginning of even better results. Other limitations include not having teacher-

specific data, witnessing numerous missing values of key variables, and having insufficient information on parental occupations, educational background and job status. Also, there is an imbalance in the size of the schools, though we believe the similar classroom sizes generate the results rather than the differential school sizes. Lastly, there are the usual caveats pertaining to measurement error in survey data, specifically in creating the self-esteem index, and the questionable nature of test scores as predictors of student achievement.

Given the public's grave concern for the educational outcomes of urban, disadvantaged youths, for particular boys, this study suggests a gender-segregated learning environment is propitious. Boys prosper and girls are at least no worse off. Future research addressing long term impacts is necessary, but it requires a different data set. Future research involving this data set will examine differences in math versus reading test scores to see where progress is greatest.

REFERENCES

- Ehrenberg, Ronald G., Daniel D. Goldhaber, and Dominic J. Brewer.** (1995) "Do Teachers' Race, Gender and Ethnicity Matter? Evidence From the National Educational Longitudinal Study of 1988." Industrial and Labor Relations Review 48 (3): 547-561.
- Evaluating Growth Using Norm-Referenced Test Scores.** (2006) Monterey: CTB/McGraw-Hill.
- Gill, Brian, et al.** (2007) "State Takeover, School Restructuring, Private Management, and Student Achievement in Philadelphia." Santa Monica: RAND Corp.
- Gilson, Judith.** (2002) "Single-Gender or Coeducation for Middle-School Girls: Does It Make a Difference in Math?" Gender in Policy and Practice: Perspectives on Single-Sex and Coeducational Schooling. Ed. Amanda Datnow and Lea Hubbard. New York: RoutledgeFalmer,,: 227-241.
- Jimenez , Emmanuel and Marlaine E.Lockheed.** (1988) "The Relative Effectiveness of Single-Sex and Coeducational Schools in Thailand." World Bank: Population and Human Resources Department, Working Paper 29, August.
- Lee, Valerie E., and Anthony S. Bry.** (1986). "Effects of Single-Sex Secondary Schools on Student Achievement and Attitudes." Journal of Educational Psychology 78 (5): 381-395.
- Lee, Valerie E., and Helen M. Marks.** (1990) "Sustained Effects of Single-Sex Secondary School Experience on Student Achievement and Attitudes." Journal of Educational Psychology 82(3): 378-392.
- Lepore, Paul C., and John R. Warren.** (1997) "A Comparison of Single-Sex and Coeducational Catholic Secondary Schooling: Evidence from the National Educational Longitudinal Study of 1988." American Educational Research Journal 34(3): 485-511.
- "Partnership Schools."** The School District of Philadelphia: Building a Better District. 9 Apr. 2007. <http://www.phila.k12.pa.us/src/partnership_schools.html>.
- "Pennsylvania School Test Scores."** The Philadelphia Inquirer 2 Mar. 2003: P, 9-10.
- School District of Philadelphia – School Profile** <http://sdp-webprod.phila.k12.pa.us/School_profiles/servlet.html> May 20, 2009.
- Porter, Judith R. & Robert E. Washington.** (1979) "Black Identity and Self-Esteem: A Review of Studies of Black Self-Concept, 1968-1978." Annual Review of Sociology 5: 53-74.
- Ransome, Whitney and Meg Milne Moulton.** (2001) "Why Girls' Schools? The Difference in Girl-Centered Education." Fordham Urban Law Journal 29: 589-599.
- Riordan, Cornelius.** (1985) "Public and catholic Schooling: The Effects of Gender Context Policy." American Journal of Education 93(4): 518-540.

- Riordan, Cornelius.** (2002) "What Do We Know about the Effects of Single-Sex Schools in the Private Sector?: Implications for Public Schools." Gender in Policy and Practice: Perspectives on Single-Sex and Coeducational Schooling. Ed. Amanda Datnow and Lea Hubbard. New York: RoutledgeFalmer: 10-30.
- Rosenberg, Morris.** (1989) Society and the Adolescent Self-Image. Revised edition. Middletown, CT: Wesleyan University Press.
- Singh, Kusum, Claire Vaught, and Ethel W. Mitchell.** (1998) "Single-Sex Classes and Academic Achievement in Two Inner-City Schools." The Journal of Negro Education 67(2): 157-167.
- Streitmatter, Janice.** (2002) "Perceptions of a Single-Sex Class Experience: Females and Males See It Differently." Gender in Policy and Practice: Perspectives on Single-Sex and Coeducational Schooling. Ed. Amanda Datnow and Lea Hubbard. New York: RoutledgeFalmer: 212-226.
- Thompson, Jennifer S.** (2003) "The Effect of Single-Sex Secondary Schooling on Women's Choice of College Major." Sociological Perspectives 46(2): 257-278.
- Tyack, David & Elisabeth Hansot.** (1990) Learning Together: A History of Coeducation in American Schools. New Haven: Yale University Press.
- Woodall, Martha.** (2006) U.S. Eases Rules on Single-Sex Schools. Philadelphia Inquirer, Philadelphia, PA, October 25:B4 C.
- Woolridge, Jeffrey M.** (2009) Introductory Econometrics: A Modern Approach 4e Thomson, South-Western: Mason, Ohio