Journal of Research in Music Education

http://jrm.sagepub.com/

Is It the Music or Is It Selection Bias? A Nationwide Analysis of Music and Non-Music Students' SAT Scores

Kenneth Elpus

Journal of Research in Music Education published online 21 May 2013
DOI: 10.1177/0022429413485601

Article

National Association for Music Education

Journal of Research in Music Education
XX(X) 1–20

DOI: 10.1177/0022429413485601
jrme.sagepub.com
SAGE

Is It the Music or Is It Selection Bias? A Nationwide Analysis of Music and Non-Music Students' SAT Scores

Kenneth Elpus[1]

Abstract

This study examined the college entrance examination scores of music and non-music students in the United States, drawing data from the restricted-use data set of the Education Longitudinal Study of 2002 (ELS), a nationally representative education study (*N* = 15,630) conducted by the National Center for Education Statistics. Analyses of high school transcript data from ELS showed that 1.127 million students (36.38% of the U.S. class of 2004) graduated high school having earned at least one course credit in music. Fixed-effects regression procedures were used to compare standardized test scores of these music students to their non-music peers while controlling for variables from the domains of demography, prior academic achievement, time use, and attitudes toward school. Results indicated that music students did not outperform non-music students on the SAT once these systematic differences had been statistically controlled. The obtained pattern of results remained consistent and robust through internal replications with another standardized math test and when disaggregating music students by type of music studied.

Keywords

music students, SAT scores, standardized tests, academic achievement

When called upon by educational policy makers to justify or support the expenditure of public funds for music programs in the schools, music educators often relay to parents, administrators, and school boards that formal music study is associated with

[1]University of Maryland, College Park, MD

Corresponding Author:
Kenneth Elpus, Assistant Professor of Music Education, University of Maryland, 2110 Clarice Smith Performing Arts Center, College Park, MD 20742, USA.
Email: elpus@umd.edu

increased performance on standardized tests. The most commonly cited statistic involves comparing the scores of music and non-music students on the College Board SAT (Americans for the Arts, 2011), one of the two primary standardized college entrance examinations offered to secondary students in the United States. The National Association for Music Education (NAfME) mentions the outperformance of music students on the SAT on many of the advocacy fliers, publications, and resources it provides to members and non-members on the "Advocacy Groundswell" section of its website (NAfME, 2011a). The strategic value of emphasizing possible links between music and SAT scores is understood easily, because improved standardized test scores are often a goal of school administrators and educational policy makers. *Newsweek* and other such publications factor SAT scores into their national rankings of the best American high schools (Newsweek, 2011), and high schools' collective college acceptance rates are coupled closely with SAT scores.

Prior research results have indicated that school music participation may be associated positively with academic achievement as measured by standardized test scores. An entire special issue of the *Journal of Aesthetic Education* (*JAE*) in 2000, titled "The Arts and Academic Achievement: What the Evidence Shows," was dedicated to examining the academic performance of arts and non-arts students. In that volume, Winner and Cooper (2000) meta-analyzed some 31 published and unpublished studies, yielding 66 separate effect sizes examining the general research question of whether arts education, broadly defined, positively influenced academic achievement. Results of the meta-analysis showed that arts education was moderately positively associated with higher achievement in math, verbal, and composite math-verbal outcomes. In the same journal issue, Vaughan and Winner (2000) sought to analyze the link between arts course work and SAT scores specifically. Using data from 12 years of national SAT means reported by the College Board in the annual *Profiles of College Bound Seniors* report, Vaughan and Winner found that students who self-reported on the SAT's Student Descriptive Questionnaire that they had pursued arts course work outscored students who reported they had not taken any arts course work. Meta-analyses of music students' performance on verbal (Butzlaff, 2000) and mathematical (Vaughan, 2000) standardized tests were somewhat inconclusive: Although positive associations were found in the correlational research literature, meta-analyses of results from the few experimental studies located in the literature showed little to no influence of music on verbal or math test scores.

Since the publication of the *JAE* meta-analyses, new research has been published examining associations between music course work and standardized test scores. Babo (2004) found that instrumental music course enrollment was associated positively with standardized test scores on the New Jersey Grade Eight Proficiency Assessment in a suburban district; however, as he added covariates to the model, the influence of instrumental music enrollment was reduced to practical and statistical non-significance. Johnson and Memmott (2006) sought to determine whether standardized test scores varied as a function of the rated quality of the school music program for elementary and middle school students; they concluded that students of higher quality music programs tended to earn higher average standardized test scores, but the effect sizes were small. In a study of public high school students in one urban district, Fitzpatrick (2006) found that instrumental music students earned consistently and statistically significantly higher test scores than did their non-instrumental peers within matched free or reduced-price lunch statuses. Fitzpatrick noted, however, that this advantage also was present in younger grades, before the instrumental music students began studying instrumental music. Kinney (2008) extended the Fitzpatrick study by adding additional controls for student mobility and family composition. He found that band students tended to score higher than did non-band students on standardized tests; similar to Fitzpatrick, Kinney found that an achievement advantage in favor of the band students existed prior to their initial enrollment in band. Helmrich (2010) found small (partial $\eta^2 = .02$) but statistically significant advantages for choral and instrumental music students on the ninth-grade Maryland Algebra/Data Analysis High School Assessment when compared to their non-music peers. Most recently, Baker (2011) found that fifth- through eighth-grade music students earned higher mean scores in English/language arts and math components of the Louisiana Educational Assessment Program than did their non-music counterparts.

Outside of the United States, similar studies of music students' achievement on standardized tests have been conducted. In British Columbia, Canada (Gouzouasis, Guhn, & Kishor, 2007), results of an observational study indicated an association between music enrollment and higher subject-area standardized test scores among high school students. The results of a randomized experiment in Montreal, Canada, showed no effects of piano instruction on subject-area standardized tests among elementary school children from low socioeconomic backgrounds (Costa-Giomi, 2004).

Several researchers have sought to determine whether the associations between music enrollment or participation and standardized test scores found in smaller studies were replicable in large observational studies using nationally representative educational data from the U.S. Department of Education's National Center for Education Statistics (NCES). Catterall (1997, 2009) determined that students of the arts represented in the National Education Longitudinal Study of 1988 (NELS) public-use data set sample were more commonly in the top quartiles of standardized test scores than were their low- or no-arts peers, although due to the methods employed, these results may not generalize beyond the NELS sample. In a more sophisticated multilevel analysis accounting for the nested structure of the NELS public-use data, Miksza (2007) found that students who self-reported music participation in the 8th and 12th grades outscored non-music students on NELS-specific standardized tests in the 8th grade and showed a greater rate of change when retested in the 12th grade. Using NELS data for adolescents and data from the Early Childhood Longitudinal Study–Kindergarten cohort for children, and controlling for various background demographics, Southgate and Roscigno (2009) found that studying music in school was associated with a small but statistically significant gain in standardized math test scores among children but not adolescents. Similarly small, yet statistically significant, effect sizes for the association of music and reading test scores were found for both children and adolescents.

Although empirical evidence supporting a modest positive association between music participation and standardized test scores seems to be mounting, two main concerns arise when interpreting the results of the past literature in this area. The first is the familiar qualification raised by arts education scholars that causal interpretation of the studies on music and test scores generally is not warranted, often due to limitations inherent in the research methodologies employed (see, for example, Eisner, 1998; Reimer, 1999, 2005). The extant research evidence, with the exception of a single randomized experiment (Costa-Giomi, 1999, 2004), derive from correlational studies that are capable of demonstrating an association between music and higher test scores but are unable to establish either the causal direction of the association or whether the association itself is an "epiphenomenon," due to one or more unmeasured variables (Winner & Cooper, 2000).

In recent empirical work, researchers have shown that there were significant preexisting, systematic differences between those students who did and did not elect high school music course work with respect to gender, race/ethnicity, socioeconomic status (SES), native language, parental education, and academic achievement prior to music enrollment (Elpus & Abril, 2011; Kinney, 2010). The systematic differences documented in the literature to date suggest that selection into elective music courses is a complex process involving many variables at many levels. If any of these variables are correlated with both the likelihood of becoming a music student and with higher test scores, then analyses comparing music to non-music students without control for the confounds likely have resulted in biased estimates of the magnitude and direction of the associations. As it happens, many of the variables that have been shown to be associated with selection into music also have been found to be associated with standardized test scores: gender (Hedges & Nowell, 1995), race/ethnicity (Card & Rothstein, 2007; Freedle, 2003), SES (Zwick & Greif Green, 2007), and native language (Abella, Urrutia, & Shneyderman, 2005). This raises the possibility that many results reported in the music and test score literature suffer from omitted-variables bias that masks the true degree and nature of the observed association between music enrollment and standardized test scores (Wooldridge, 2002). The only effect size in the literature resulting from a randomized experiment (Costa-Giomi, 2004) was small for reading and nil for math, suggesting that the results from uncontrolled or inadequately controlled observational studies may be biased upward. There is a need for further research in this area using data that not only include measures of music enrollment and test scores but also provide a rich set of covariates to help reduce selection bias.

Purpose of the Study

The choice of enrolling in music during high school seems to be the result of a complex selection process—not yet fully understood—that leads to substantial population differences between students who do and do not elect music that are measurable prior to the commencement of music course work. Research methodologists have shown empirically that observational studies can, with careful design and under certain conditions, substantially reduce bias due to selection in analytical results (Steiner, Cook, Shadish, & Clark, 2010). Steiner and colleagues demonstrated that selection bias could be reduced most effectively with either a complete model of selection *or* the appropriate choice of multiple covariates from multiple heterogeneous domains. The purpose of this study was to compare the SAT scores of a nationally representative sample of music and non-music students, while controlling for a series of preexisting systematic differences between the groups that have been demonstrated in prior research, based on an estimate of nationwide enrollment rates for music using transcript data.

In the primary analysis for this study, I examined the college entrance exam scores of music students and non-music students in the nationally representative, multiwave Education Longitudinal Study of 2002 (ELS). I classified "music" and "non-music" students in the ELS sample using the sample members' complete high school transcripts, which are available in the restricted-use ELS data set. Because not all students in the United States choose to take a college entrance exam, I conducted additional analyses to expand the analytic sample beyond those ELS sample members who had taken a college entrance exam by substituting a standardized test unique to ELS that was administered to the entire sample at each wave of in-school data collection. Finally, I attempted to discern whether the temporal ordering of initial music enrollment between a pre- and post-test provided support for a causal effect of music enrollment on increased test scores.

Data, Research Design, and Methods

Data

For the present study, I used the restricted-use data set from ELS, an ongoing multiwave longitudinal study of a nationally representative cohort of students who were sophomores in U.S. public and private high schools in 2002 (Ingels et al., 2007). Sample members were resurveyed in 2004 and 2006. A third ELS follow-up began collecting data in the summer of 2012.

In the ELS data set, I was able to observe sample members' college entrance exam scores as reported by the Educational Testing Service or ACT, Inc., a continuous measure of SES,[1] race and ethnicity, family composition (single- vs. dual-parent household), academic achievement as measured by an ELS-specific standardized test, academic achievement as measured by grade point average (GPA) in each high school grade, and a host of self-reported time use and attitudinal measures related to school engagement. I also was able to observe complete high school transcripts for nearly all sample members. Whereas most previous music education research using NCES data identified music and non-music students using the self-reported answer to a question in the survey about music participation (Elpus & Abril, 2011; Miksza, 2007, 2010; Morrison, 1994), for the present study, I used sample members' official high school transcripts to identify the music students in the sample.

Analytic Sample

Data preparation for the present study comprised the selection of an appropriate analytic sample and the identification of the music and non-music students from within the analytic sample. To be included in the analytic sample, I required that ELS sample members had been survey respondents in both the base year ("BY") and first follow-up ("F1") ELS waves; additionally, they must have had complete transcript data available. These conditions were met for the great majority of ELS sample members (N = 13,530; 87% of the initial BY sample).

Music students were identified from the transcript data by examining the standardized Classification for Secondary School Courses (CSSC) coding that NCES appends to each course in the transcript data set.[2] Given that the data entry personnel employed by NCES in the coding of courses were not experts in music education, and because the CSSC codes consequently were not applied uniformly to similar music courses between different schools, I manually verified that each of the 19,920 courses coded as "music" in the data set was in fact a music course and was classified properly in its subarea (e.g., band, choir, orchestra, music theory, guitar, piano, etc.). The final analytic data set contained variables indicating music enrollment by course type, a continuous measure of the number of Carnegie units of credit earned in music, all outcome variables, and all covariates for the entire analytic sample of music and non-music students (N = 13,530).

The dependent variable of interest in the primary analysis was the college entrance exam score of music and non-music students in the analytic sample. In the United States, those students who take college entrance exams most often take either the College Board SAT, administered by the Educational Testing Service, or the ACT, administered by ACT, Inc. The choice of which exam to take is often a function of geography (the ACT is traditionally more common in the Midwest, and the SAT is more common on the coasts) and personal content preference (Lewin, 2005), but colleges view the exams as interchangeable and place standardized test scores of all applicants on the same scale using score concordance tables jointly published by the two testing agencies (ACT Inc. & the College Board, 2006). For the purposes of this study, I used sample members' college entrance exam score as expressed on the SAT scale regardless of which exam the student had taken. Not all students in the United States choose to take the SAT or the ACT; as such, a college entrance exam score is observable only for 9,530 ELS sample members. Of these, 9,110 were BY and F1 respondents with complete transcript data, and 8,800 had survey data available for the covariates (56% of the BY sample). Given the emphasis on college entrance exam scores in the prior literature, I limited the analytic sample for the primary analysis to those who had taken college entrance exams; this restricts the generalizability of this analysis to that self-selected population.

A broader question of interest is whether there is a link between music enrollment and a more general notion of standardized test scores not focused solely on college entrance exams. To answer this broader question in the second analysis, I substituted an ELS-specific standardized test in math that was administered to nearly all sample members in both the BY and F1 waves of the study. This expanded the analytic sample to 13,390 with an observable test score, 12,840 of whom had survey data available for the covariates (82% of the BY sample). This second analysis increases the generalizability of the results to include students who were not college bound or who planned to attend colleges for which standardized test scores were optional.

Covariates

School Level. I controlled for between-school variation by including school-level fixed effects in the analysis (Wooldridge, 2002). Using school fixed effects accounted for the multilevel or "nested" structure of the data by allowing each school to have its own intercept in the regression. This had the net effect of removing all between-school variance from the error term, essentially controlling for all observable and unobservable school differences that might have influenced the observed test scores. Removing this variance from the error term also increases the precision of the point estimates; however, the trade-off for using fixed effects to remove all between-school variation is the inability to select strategically school-level covariates based on theory.

Individual Level. Following the suggestion of Steiner and colleagues (2010), I used a vector of individual covariates selected from multiple, heterogeneous domains that theoretically or empirically are linked to the outcome measure in an effort to reduce selection bias. Prior research has shown that the covariates I employed are related empirically to selection into music or to systematic differences between the music students and non-music students. Table A1, available in the online appendix to this article (see online supplemental material available at http://jrme.sagepub.com/supplemental), presents the covariates chosen for this analysis, sources for the data, and citations to past research establishing their relevance.

In the analyses, some covariates were recoded from their encoding in the ELS data set. Racial and ethnic categorizations were dichotomized to "White" and "non-White." I dichotomized race and ethnicity to follow methodological choices of previous researchers (Stewart, 1991) and to ease the interpretation of the model. Family composition was dichotomized to "dual parent/guardian family" and "single parent/guardian family," following the methodology of Kinney (2010) and Elpus and Abril (2011). The variable for college degree expectation was dichotomized to "expects bachelor's degree or higher" and "does not expect bachelor's degree" to ease the interpretation of the model.

Empirical Approach

Theoretical Models. The relationship between college entrance exam scores and music enrollment as conceptualized in this study is modeled in Equation 1:

$$SAT_{ij} = \alpha_j + \beta Music_{ij} + \gamma_{1 \ldots k} Covariates_{ij} + \varepsilon_{ij} \quad (1)$$

where SAT_{ij} is the standardized test score for the ith student in the jth school, α_j is the school fixed effect for school j, $Music_{ij}$ is either a dichotomous indicator of music enrollment for the ith student in the jth school *or* a continuous measure of the number of credits earned in music, $Covariates_{ij}$ represents the vector of covariates from Table A1 for the ith student in the jth school, and ε_{ij} is the error term. The parameter of interest, β, represents the effect of music enrollment on college entrance exam scores net of the covariates. While instructive, this model does not illuminate possible differential effects among the *types* of music studied. To address this deficiency, I also estimated a modified version of Equation (1) using seven disaggregated indicators of music subarea, as modeled in Equation (2):

$$SAT_{ij} = \alpha_j + \beta_{1\ldots7}TypeOfMusicStudied_{ij} + \gamma_{1\ldots k}Covariates_{ij} + \varepsilon_{ij} \tag{2}$$

where the seven βs represent the effects of each type of music studied on the test score.

One important element of establishing potential causality is the notion of temporal ordering; a proposed cause must precede a proposed effect in time (Shadish, Cook, & Campbell, 2002). The models represented in Equations 1 and 2 consider the ELS cohort across the 4 years of high school as a whole and include music courses that may have occurred at any point in the high school career. Because the BY ELS-specific standardized test was administered in the 10th grade, and most students take college entrance exams in the spring of 11th and fall of 12th grades, I was able to construct a model with proper temporal ordering (Pretest 10th grade – Music Course(s) 11th grade – SAT/ACT late 11th grade/early 12th grade) if I restricted the analytic sample to exclude those students who took music in the 9th or 10th grade and classified as music students only those students who *began* high school music course work in the 11th grade. Though this subset of late starters may not be representative of the broader population of music students, the advantage in analyzing the data under this restriction is that course work in music here is conceptualized as an "intervention" between a pre- and post-test in the quasi-experimental sense of the term, as modeled in Equation 3:

$$SAT_{ij} = \alpha_j + \beta Music11_{ij} + \gamma Pretest10_{ij} + \delta_{1\ldots k}Covariates_{ij} + \varepsilon_{ij} \tag{3}$$

where $Music11_{ij}$ is a dichotomous indicator for having first enrolled in music during the 11th grade and $Pretest10_{ij}$ is the BY ELS standardized test score. The gain in proper temporal ordering for this final analysis is a trade-off earned with the concomitant reduction in analytic sample size (for this analysis, N = 5,980) and the potential that the students identified as "music students" in this analysis were subject to selection biases that may differ systematically from those for the full music student sample.

Estimation. I estimated parameters for Equations 1, 2, and 3 using fixed-effects ordinary least squares (OLS) regression. The analyses were weighted using the probability weights provided by NCES for the BY-to-F1 panel cohort to ensure estimates obtained were nationally representative and accounted for oversampling of minority students. In estimating the parameters, I used cluster-robust adjusted standard errors (Rogers, 1993), which corrects significance tests for the design effects introduced by the ELS school-clustered sampling. I estimated each model sequentially, starting with an estimate excluding all covariates and then adding covariates successively in blocks by domain to determine whether the uncontrolled models suffer from selection bias. Given the large sample sizes in the analyses, I set a conservative .01 alpha level for significance testing.

Missing Data Strategy. For all covariates except demographics and BY test scores, a set of missing data dichotomy variables (Cohen, Cohen, West, & Aiken, 2003) is used to maintain the analytic sample size and to preserve representativeness in the sample.[3] Sample members with missing demographic or BY test score information were generally survey non-respondents; I decided to delete these cases from the analyses.

Results

National Estimates of Transcript-Indicated Music Course Enrollment

Tables 1, A2, and A3 present the national estimates of the number of music students in the U.S. senior class of 2004 as indicated by their high school transcripts. As population estimates are based on complex survey sampling, the descriptive statistics are reported here with a standard error and a corresponding 95% confidence interval. Because these tables are derived from the transcript data, I used the cross-sectional transcript weight to ensure the estimates generalize to the entire cohort of students who were sophomores in 2002 and seniors in 2004. As seen in Table 1, approximately 36.38% (SE = 0.83%, 95% CI [34.76%, 38.01%]) of students in U.S. high schools—an estimated 1.127 million in the 12th-grade cohort of 2004—earned at least one credit of music on their official transcript by the time they finished high school. Students who took any music in high school earned a mean of 2.25 credits in music (SE = 0.032, 95% CI [2.18, 2.31]) on their transcripts by graduation.

Table A2, available in the online appendix to this article (http://jrme.sagepub.com/supplemental), presents subpopulation estimates based on music persistence throughout high school. The estimates reported in Table A2 do not double count the simultaneous pursuit of multiple music courses; students are considered as having taken multiple years of music only if one or more music courses were taken in each of the grade levels from 9 through 12, although these may have been in any combination. For example, a student who enrolled in music for 9th and 11th grades would be considered in the 2+ years estimate, but a student who enrolled in band and choir simultaneously during 9th grade and took no more music would be considered only in the 1+ years estimate. As displayed in Table A2, over 1 out of every 3 students earned at least one credit in music, and nearly 1 in 10 U.S. high school students graduated having taken 4 years' worth of music courses in high school. Table A3 (http://jrme.sagepub.com/supplemental) presents subpopulation estimates based on the grade level in which students were enrolled in their music courses.

Descriptive Statistics for Covariates

For each covariate used in the analyses, I calculated the variable's mean (for continuous measures) and proportion (for dichotomous measures) across the entire population and within the subpopulations of music and non-music students. As an empirical check of the covariate's potential relationship to selection bias, I calculated t statistics for mean or proportional differences of each covariate between the music student and

Table 1. Estimates of the National Music Student Population.

Course Taken	Subpopulation Estimate	Linearized *SE* (in thousands)	95% Confidence Interval (in thousands)	Percentage of Population
Any music	1,127,000	32.80	[1,062, 1,191]	36.38
Choir	538,000	22.26	[494, 581]	17.37
Band	416,000	16.25	[384, 448]	13.44
Orchestra	67,000	6.51	[54, 80]	2.15
Music Theory	68,000	7.90	[52, 83]	2.19
Guitar	65,000	7.07	[51, 79]	2.10
Piano	47,000	6.57	[35, 60]	1.53
Secondary General	190,000	18.62	[153, 226]	6.13

Note: The total student population estimate (N = 3,496,440) and all subpopulation estimates are based on the transcript sampling weight. Design df = 390 for all estimates. Subpopulations are the number of students with at least one credit-earning year of each type of music appearing on their transcript. Totals of subpopulations do not equal the "any music" estimate due to rounding and the possibility of one student enrolling in multiple types of music.

non-music student subpopulations. In the entire population, 63.4% of students were White; 67.6% of music students were White and 60.94% of non-music students were White. The proportion difference between music and non-music students was significant, $t(390) = 4.48$, $p < .001$. Among all students nationally, 12.5% received special education services under an Individualized Education Plan (IEP). Among music students, 9.25% received special education services under an IEP; for non-music students, the percentage was 14.37%. The proportion difference between music and non-music students was significant, $t(390) = 5.04$, $p < .001$. Table A4, available in the online appendix to this article (http://jrme.sagepub.com/supplemental), presents the population and subpopulation means for the continuous covariates as well as the t statistics and p values for mean differences between music and non-music students. As expected from the prior literature, all mean differences between music and non-music students on the covariates were statistically significant.

Estimates of the Effect of Music Enrollment on SAT Scores

Using a Dichotomous Indicator of Music Enrollment. Table A5 (http://jrme.sagepub.com/supplemental) displays the results of regressing SAT scores on high school music enrollment coded dichotomously; the variable is set to one for students who earned credit in at least one music course at any point in high school and encoded zero otherwise. Model 1, with no covariates to reduce selection bias, illustrates the simple mean difference between these two groups; this initial model shows that the music student mean SAT score is greater than the non-music student mean SAT score by 37.30 points—more than one sixth of a standard deviation. However, the robustness of the association, both in terms of practical significance and in statistical significance, is diminished quickly as additional covariates are added to the regression, suggesting that the estimate for the effect of music enrollment in the uncontrolled model is biased upward. Adding school fixed effects, which represent the relative advantage or disadvantage to SAT scores for attending any particular school, diminishes the coefficient on music enrollment modestly and slightly increases the precision of the estimate. The coefficient on music enrollment is reduced considerably when demographic variables are added to the regression. The coefficient declines 31%, to 25.61, when this domain of covariates is held constant in Model 3. Model 4 adds the covariates for prior academic achievement and IEP status; holding these variables constant eliminates both the practical and statistical significance of the coefficient for music enrollment, which falls to 0.23. In Model 4, the coefficient on the indicator for music enrollment is statistically indistinguishable from zero. This is perhaps the clearest evidence yet that an uncontrolled regression of SAT scores on music enrollment suffers from omitted-variables bias. This analysis suggests that the often-demonstrated correlation between music enrollment and SAT scores is more likely a function of music students' relative advantages in SES and prior academic achievement and music students' relatively lower incidence of being classified as students with special needs.

Although adding covariates in Models 5 and 6 for time use and school attitudes alters the sign on the coefficient for music enrollment from positive to negative, the estimate remains statistically non-significant. Additionally, in Models 5 and 6, there remains the possibility that covariates in the time use and school attitudes domains are endogenous with regard to music enrollment—being involved in music, for example, may reduce the time available for video games and TV after school. Accordingly, the negative sign should not be interpreted necessarily as a disadvantage for music students. Model 4, showing no significant advantage for music students in SAT score once prior academics, IEP status, and demographics are controlled, is likely the most trustworthy estimate of the influence of music enrollment on SAT scores presented. Enrollment in music courses appears to have made no difference in SAT scores.

Music Enrollment "Dosage" and SAT Scores. In the previous analysis, music enrollment was operationalized as a simple binary. Because I was able to observe the amount of music course work (in Carnegie units) earned by music students on their high school transcripts, I re-estimated the models using credits earned as a continuous measure of sustained music enrollment. This analysis estimates the effect of each additional year of music enrollment on SAT scores and is predicated on the notion that more sustained music enrollment may represent more active musical engagement. Table A6 (http://jrme.sagepub.com/supplemental) displays the results of the music enrollment dosage estimates. The uncontrolled Model 1 suggests a 15.3-point gain in SAT scores for each additional yearlong music course taken. Adding school fixed effects to the regression, as in Model 2, reduced the advantage to 12.82 points per year of music. However, once demographics (Model 3) and prior academic achievement (Model 4) were entered into the regression, there remained no practical or statistical significance on the coefficient for music enrollment dosage. Applying a less conservative standard for statistical significance than the α value of .01 used in this study suggests that students who persist in music for 4 years of high school have only an 8.68-point advantage when compared to students who earn no credits in music. This represents 0.04 of a standard deviation in SAT scores—in other words, a negligible score gain in practical terms. The practical negligibility of this gap is put into appropriate perspective when considered in context of the confidence intervals for SAT scores reported by the College Board for use in interpreting SAT scores: The College Board (2011) indicates that any one reported SAT score on the 400-to-1600 scale is interpreted more appropriately as the center of a ±60- to 80-point range (0.30 to 0.40 of a standard deviation), which most likely contains the student's true score.

Disaggregating Type of Music Studied. In the previous analyses, I considered all the possible subareas of music as equivalent. It is possible, however, that different kinds of music enrollment may influence SAT scores differentially—perhaps there is a differential link for instrumental versus vocal performance enrollment or for music theory versus music performance enrollment, for example. Table A7 (http://jrme.sagepub.com/supplemental) presents the estimates for the disaggregated analysis. In this table, coefficients reported estimate the unique influence of each kind of music enrollment on SAT scores against a non-music student comparison. The general pattern of results found for the aggregated music enrollment model holds in the disaggregated model. Intriguingly, the initial, uncontrolled model suggests that the SAT score advantages earned by instrumental (i.e., band and orchestra) students may be driving the observed aggregated bivariate relationship between music enrollment and SAT scores. The coefficients for other types of music were not statistically significant, with the notable exception of a significant 10-point SAT *deficit* for choral students versus non-music students in Model 4 when school fixed effects, demographics, prior academic achievement, and IEP status are controlled. As in the previously reported results, the most robust predictors of SAT score remain SES, prior academic achievement, and IEP status. The results of this analysis suggest not only that selection into music drives the observed bivariate relationship between music enrollment in the models with no selection control, but that selection into instrumental music may be correlated most highly with the omitted variables of the spurious music enrollment–SAT association. This raises the likelihood that selection into music is not only a complex process, but that selection into various types of music may be differentiated.

Internal Replication With a Larger Analytic Sample. There remains the possibility that a link between music enrollment and standardized test scores may exist in the broader population of U.S. high school students that somehow is masked by differential selection into SAT or ACT test taking. To assess this possibility, I re-estimated the models substituting the ELS-specific math test that was administered to all sample members in their sophomore and senior years. In this analysis, I was able to use the sophomore year score as a true, rather than proxy, pretest covariate.

As reported in Tables A8, A9, and A10, available in the online appendix (http://jrme.sagepub.com/supplemental), the analytic sample size increased dramatically when the ELS idiosyncratic math test was used as the dependent variable in the OLS regression estimates (to $N = 12,840$ in these estimates from $N = 8,800$ for the models using SAT scores). The uncontrolled Table A8 Model 1 showed a mean difference for music and non-music students of 2.14 points on the senior-year (F1) ELS standardized math test, or roughly one fifth of a standard deviation. The advantage for music students is reduced slightly when school fixed effects are added in Model 2. Controlling for demographics in Model 3 reduces the mean difference by 22%, to 1.64. Model 4, which adds measures of prior academic achievement, eliminates the practical and statistical significance of the coefficient of the indicator for music enrollment. Dosage (Table A9) and disaggregated (Table A10) analyses similarly replicated the SAT results reported earlier. Thus, the pattern of results seen in the earlier estimates using SAT score as the outcome measure are replicated in their entirety for the larger analytic sample size. This suggests that the lack of a music enrollment effect on SAT scores demonstrated earlier holds when using a standardized math test administered to a more diverse sample as the outcome measure. Music enrollment had no effect on standardized math test scores, once demography, prior academic achievement, and IEP status were controlled.

Temporally Ordered Analysis. Table A11 (http://jrme.sagepub.com/supplemental) presents the result of the temporally ordered analysis, which conceptualizes initial music enrollment as an intervention between a pre- and post-test. Although the coefficient for music enrollment in this analysis did not achieve the .01 level of significance, the sign of the estimate indicates that music students first enrolling in music as high school juniors were at a test score *deficit* to their non-music peers ($p = .049$). Applying the looser .05 level of significance, the deficit is robust in the initial, school fixed effects, and demographically controlled Models 1, 2, and 3. However, consistent with the prior pattern of results, the negative effect is statistically non-significant—even at the looser .05 standard—in Model 4 once prior academic achievement and IEP status are added to the regression. This suggests that while academically superior students in the cohort may have been drawn to music earlier in high school, academically inferior students may have taken up music later in the high school career, and these music courses did not raise their level of academic achievement in comparison to their non-music peers.

Discussion

In the present study, I sought to determine whether music students outscored non-music students on college entrance examinations while controlling for a series of covariates related to selection into music. The analyses demonstrate that music students in the U.S. high school class of 2004 did not outperform non-music students on college entrance exams or on standardized math tests. The results remained consistent across all specifications of the models, whether aggregating or disaggregating types of music enrollment, increasing analytic sample size by substituting a standardized test given to all ELS respondents as the outcome, temporally ordering the analyses so that music enrollment intervened between a pre- and post-test, or regressing standardized test scores against a continuous measure for the number of credits earned in music.

The most robust and consistent predictors of SAT and math standardized test scores were prior academic achievement (operationalized as ninth-grade academic GPA and a BY pretest), IEP status, and SES.

Although contrary to much of the prior literature examining bivariate relationships between music and test scores (Americans for the Arts, 2011; Catterall, 1997, 2009; Catterall, Chapleau, & Iwanaga, 1999; Cheek & Smith, 1999; Fitzpatrick, 2006; Gouzouasis et al., 2007; Johnson & Memmott, 2006; Kinney, 2008; Miksza, 2007; NAfME, 2011b; Vaughan & Winner, 2000), the absence of a significant effect of music enrollment on test scores found in this study replicates the results of the only randomized experimental study of this nature reported in the music education literature (Costa-Giomi, 2004). The results obtained in this study demonstrate that correlational claims that music students outperform non-music students on the SAT or other standardized tests are perhaps untrustworthy because of omitted-variables bias. The bivariate relationships between SAT scores and music enrollment that seem to exist between music and non-music students may not be an effect of the music courses themselves. Rather, in the present study, they appear to have been a function of the selection process into music that, for reasons yet to be explicated in the literature, differentially favored those students who are already more likely to outperform others on educational measures. Although some research in music education has begun to explore the differential selection of students into elective music courses (Elpus & Abril, 2011; Kinney, 2010; Stewart, 1991), and other researchers have suggested that selection into music *might* drive observed test score differences between music and non-music students (Fitzpatrick, 2006; Kinney, 2008; Winner & Cooper, 2000), the present study is the first to test specifically, using nationally representative data, the hypothesis that observed score advantages for music students are due not to music courses but to selection bias.

Importantly, there is prior work in this area by researchers who have attempted at least some statistical control for the demographic differences between music and non-music students. Miksza (2007, 2010), for example, included similar SES variables as covariates in his hierarchical linear model analyses of academic and developmental outcomes of music and non-music students using NCES data. Miksza (2007) found that SES was a more robust predictor of 8th-grade standardized test scores and of the rate of change in standardized test scores from 8th through 12th grade than was music participation. Music participants in Miksza's (2007) analysis began the 1988 NELS with higher standardized test scores than their non-music peers, but the gap remained relatively stable. Higher SES students, on the other hand, began with higher scores than low-SES students, and the score gap between them increased over time. Although a different analytic technique was employed in the present study, the results of the present investigation support and extend Miksza's (2007) conclusions. Not only did music students begin with higher levels of SES and prior academic achievement, but it was these systematic differences—not an effect of music itself—that led to the eventual differences observed in 12th grade.

The present study also demonstrates that controlling for demographic factors is "not enough" to reduce omitted-variables bias in comparisons of music and non-music students. The models estimated here that controlled only for demographics produced results that were biased upwardly when compared to models that controlled for demographics and prior academic achievement. The finding that demographic covariates do not eliminate all the bias in estimates obtained by observational studies comparing music and non-music students is in line with the results of within-study comparisons examining covariate choice in the quasi-experimental educational research methodology literature (Steiner et al., 2010). At present, differential selection into elective music remains poorly understood. Replications of within-study comparisons, similar to the work of Steiner and colleagues (2010), within our own field would help illuminate the mechanisms of selection into music. Understanding the selection process would guide future music education quasi-experimental and observational research better in the appropriate choice of covariates. Beyond improved quasi-experimental practice, additional randomized controlled trials, similar to the extant Costa-Giomi (1999, 2004) study, would considerably improve the base of research evidence comparing the educational outcomes of music and non-music students, given that properly executed randomized experiments are theoretically free of selection bias.

The "dosage" analysis in this study is also important to consider in terms of understanding selection into music. A philosophical, as well as empirical, question arises when classifying music and non-music students through administrative transcript data that does not arise when using self-report survey questions about music participation. Although self-report data sometimes are thought to be less inherently trustworthy than administrative data, a self-reported question about music participation allows the respondent to consider his or her own musical involvement and determine whether it is substantial enough to warrant reporting. In analyzing administrative data, the analyst is left to decide "how much music" is required to classify a student as a music student. Although some may argue that the one course criterion employed in the dichotomous classification of music enrollment status employed in this study is not restrictive enough to separate "real" music students from dilettantes or those participating in a required music appreciation course, the dosage study measures sustained musical involvement through high school. The dosage measure even accounts for students who doubled up and had more than 4 years' worth of music courses appearing on their transcripts. The results presented here indicate that even those students with sustained music enrollment did not outscore their non-music peers on standardized tests once school fixed effects, demographics, and prior academic achievement are taken into account. This again suggests that the selection process into sustained music enrollment likely is bound up in these factors, which also predispose students to higher standardized test scores.

The disaggregated analysis also points to the complexity of the music selection process. The initial bivariate relationships between the type of music course work pursued and the standardized test scores suggest that instrumental music students outperform choral students and non-performance students. Indeed, the results of the present study suggest that it is the outperformance of band and orchestra students in the uncontrolled models that drives the entire observed bivariate advantage for music students' standardized test scores. That choral students in the controlled models significantly underperformed the non-music students is intriguing to consider and echoes some of the results obtained by Johnson and Memmott (2006). Perhaps differential selection processes drive enrollment into different types of music courses—the cost of an instrument and private lessons, perhaps, might manifest in that more economically advantaged students are selected into instrumental classes and this increased level of economic advantage is associated with higher test scores (Zwick & Greif Green, 2007). Other possibilities for the performance discrepancy exist, and there is a need for further research on both the academic achievement–music subarea relationships and the potentially differing selection processes for each music subarea.

Winner and Cooper (2000) suggested that, at the time of their writing, there was no evidence "yet" for a causal link between music participation and greater academic achievement. The present study provides at least some evidence opposing the existence of a causal link between high school music course taking and standardized test scores. As commonly construed in social science research, causation requires the satisfaction of three conditions: (a) that the cause always covaries with the effect, (b) that the cause always precedes the effect in time, and (c) that all alternative explanations for the effect can be ruled out (Shadish et al., 2002). Evidence from prior research (Vaughan & Winner, 2000) generally has established the first of these conditions for music participation and SAT scores, albeit with the caveat that the research that establishes this condition is correlational and likely suffers from omitted-variables bias. The present study contradicts the first condition: If one views the design of this investigation as a well-controlled correlational study, then its results suggest that the covariation between music enrollment and test scores does not hold under well-controlled conditions. The present study also contradicts the second condition: that the cause always must precede the effect in time. The temporally ordered analysis reported here shows that music enrollment does not always precede a higher test score. In that analysis, students who took pretests in 10th grade, began their first high school music courses in 11th grade, and then took an SAT at the end of 11th grade or a math standardized test in the 12th grade scored lower on SATs and on the math standardized test than did their peers who had never taken music. The music student deficit in SAT and standardized test scores in this temporally ordered analysis, although statistically nonsignificant, stands in contrast to the notion that music enrollment might "cause" increased test scores (e.g., Catterall, 2009; Helmrich, 2010), at least not for students who first took up music course work in the second half of the high school career. Prior correlational research notwithstanding, the results of this study extend Winner and Cooper's (2000) conclusions in suggesting not only that there is no evidence "yet" for a causal link but that a causal link between music and increased test scores may not exist. Certainly, if the effect of music enrollment on SAT scores was robust, one might reasonably expect positive effects in the temporally ordered analysis. A plausible explanation for the results reported here, however, is bound up in the complexity of the theoretical selection process into music. While academically superior students may be drawn to music courses beginning in the 9th- or 10th-grade years, it seems possible from this analysis that academically inferior students may take up music course work later in the high school career. These students may be attempting to fill an arts requirement as upperclassmen, may be pursuing a schedule perceived to be less academically rigorous as they complete requirements for other course work, or may require remedial work in the 9th and 10th grades that precludes earlier enrollment in music.

Practically, it may be time to question seriously, or even retire, the advocacy argument that implies music students are more academically successful on standardized tests than their non-music peers because of their enrollment in music. A more candid appraisal of the current body of research literature might suggest that music is somehow attractive to those students who are already more likely to perform well academically and, as such, may serve as an important artistic outlet with positive developmental benefits for those students who choose to study it. This argument, although possibly less convincing than the common SAT claim, is more concerned with the intrinsic, rather than the extrinsic, value of music and arts education (as in Eisner, 2002). A more pressing concern for the profession, however, is to begin to understand the root cause of the preexisting systematic differences between music and non-music students. The extant research offers no concrete examination of why students who are more predisposed toward academic success are enrolled in music in greater numbers than are their peers, and there is little extant theory to guide an empirical investigation into this phenomenon.

Beyond the SAT score comparisons, the present study is also the first to document the national uptake rates of music courses using the high school transcripts of ELS sample members. Whereas the self-reported measure of senior-year participation in music ensembles puts the participation rate at 21% (Elpus & Abril, 2011), this examination of high school transcripts across all 4 years in high school and including all possible types of music courses shows that 36.38% of students nationwide graduate high school having earned at least 1 year's worth of credit in music, with the large ensembles being the most popular forms of music classes undertaken by the U.S. class of 2004. This estimate represents a slight increase in music enrollment rates from the last nationwide transcript analysis appearing in the music education literature (Stewart, 1991), which reported that 35.7% of students in the U.S. class of 1982 had been enrolled in music. That participation rates would have increased between 1982 and 2004 is somewhat contrary to the current conventional wisdom in the field, and yet, these estimates represent only two data points, which cannot confirm a trend. Further research using U.S. Department of Education transcript data would help the field understand the ongoing trends in music enrollment.

Author's Note

This article is based on the author's doctoral dissertation, "Justifying Music Education: Econometric Analyses of Issues in Music Education Policy," completed at Northwestern University in June 2011.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Notes

1. The National Center for Education Statistics calculates the socioeconomic status variable using an algorithm that accounts for family income, parental educational attainment, and parental occupational prestige. The calculation is explained fully in Ingels et al. (2007).
2. The process through which the music students were identified using transcript data is briefly described here. It is detailed exhaustively in Elpus (2011).
3. I also estimated the models without using the missing data dummy variables (analyzing only complete cases, resulting in a smaller analytic sample). The pattern of results obtained was consistent with the reported results, although the specific point estimates vary. Tables of results not using this technique are available from the author.

References

Abella, R., Urrutia, J., & Shneyderman, A. (2005). An examination of the validity of English-language achievement test scores in an English language learner population. *Bilingual Research Journal*, *29*(1), 127–144. doi:10.1080/15235882.2005.10162827

ACT Inc., & the College Board. (2006). *ACT-SAT concordance tables*. New York, NY: College Board. Retrieved from http://professionals.collegeboard.com/profdownload/act-sat-concordance-tables.pdf

Americans for the Arts. (2011). *Arts students outperform non-arts students on SAT*. Washington, DC: Author. Retrieved from http://www.artsusa.org/pdf/get_involved/advocacy/research/2011/sat_artsed11.pdf

Babo, G. D. (2004). The relationship between instrumental music participation and standardized assessment achievement of middle school students. *Research Studies in Music Education*, *22*(1), 14–27. doi:10.1177/1321103x040220010301

Baker, R. A., Jr. (2011). The effects of high-stakes testing policy on arts education. *Arts Education Policy Review*, *113*(1), 17–25. doi: 10.1080/10632913.2012.626384

Butzlaff, R. (2000). Can music be used to teach reading? *Journal of Aesthetic Education, 34*(3/4), 167–178.

Card, D., & Rothstein, J. (2007). Racial segregation and the Black–White test score gap. *Journal of Public Economics, 91*(11/12), 2158–2184. doi:10.1016/j.jpubeco.2007.03.006

Catterall, J. S. (1997). *Involvement in the arts and success in secondary school*. Washington, DC: Americans for the Arts. Retrieved from http://goo.gl/JLuvd

Catterall, J. S. (2009). *Doing well and doing good by doing art: A 12-year national study of education in the visual and performing arts*. Los Angeles, CA: I-Group Books.

Catterall, J. S., Chapleau, R., & Iwanaga, J. (1999). Involvement in the arts and human development: General involvement and intensive involvement in music and theater arts. In E. B. Fiske (Ed.), *Champions of change: The impact of the arts on learning* (pp. 1–18). Washington, DC: Arts Education Partnership.

Cheek, J. M., & Smith, L. R. (1999). Music training and mathematics achievement. *Adolescence*, *34*(136), 759–761.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.

College Board. (2011). *Understanding your scores*. Retrieved from http://sat.collegeboard.com/scores/understanding-sat-scores

Costa-Giomi, E. (1999). The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education*, *47*(3), 198–212. doi:10.2307/3345779

Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music*, *32*(2), 139–152. doi:10.1177/0305735604041491

Eisner, E. W. (1998). Does experience in the arts boost academic achievement? *Art Education*, *51*(1), 7–15.

Eisner, E. W. (2002). *The arts and the creation of mind*. New Haven, CT: Yale University Press.

Elpus, K. (2011). *Justifying music education: Econometric analyses of issues in music education policy* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3456550)

Elpus, K., & Abril, C. R. (2011). High school music ensemble students in the United States: A demographic profile. *Journal of Research in Music Education*, *59*(2), 128–145. doi:10.1177/0022429411405207

Fitzpatrick, K. R. (2006). The effect of instrumental music participation and socioeconomic status on Ohio fourth-, sixth-, and ninth-grade proficiency test performance. *Journal of Research in Music Education*, *54*(1), 73–84. doi: 10.1177/002242940605400106

Freedle, R. O. (2003). Correcting the SAT's ethnic and social-class bias: A method for reestimating SAT scores. *Harvard Educational Review*, *73*(1), 1–43.

Gouzouasis, P., Guhn, M., & Kishor, N. (2007). The predictive relationship between achievement and participation in music and achievement in core grade 12 academic subjects. *Music Education Research*, *9*(1), 81–92.

Hedges, L., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*(5220), 41–45. doi:10.1126/science.7604277

Helmrich, B. H. (2010). Window of opportunity?: Adolescence, music and alegebra. *Journal of Adolescent Research*, *25*(4), 557–577. doi:10.1177/0743558410366594

Ingels, S. J., Pratt, D. J., Wilson, D., Burns, L. J., Currivan, D., Rogers, J. E., & Hubbard-Bednasz, S. (2007). *Education longitudinal study of 2002: Base year to second follow-up data file documentation* (NCES 2008-347). Washington, DC: National Center for Education Statistics.

Johnson, C. M., & Memmott, J. E. (2006). Examination of relationships between participation in school music programs of differing quality and standardized test results. *Journal of Research in Music Education*, *54*(4), 293–307. doi:10.1177/002242940605400403

Kinney, D. W. (2008). Selected demographic variables, school music participation, and achievement test scores of urban middle school students. *Journal of Research in Music Education*, *56*(2), 145–161. doi:10.1177/0022429408322530

Kinney, D. W. (2010). Selected nonmusical predictors of urban students decisions to enroll and persist in middle school band programs. *Journal of Research in Music Education*, *57*(4), 334–350. doi: 10.1177/0022429409350086

Lewin, T. (2005, March 4). Strivers sharpen no. 2's for different college test. *The New York Times*. Retrieved from http://www.nytimes.com

Miksza, P. (2007). Music participation and socioeconomic status as correlates of change: A longitudinal analysis of academic achievement. *Bulletin of the Council for Research in Music Education*, *172*, 41–58.

Miksza, P. (2010). Investigating relationships between participation in high school music ensembles and extra-musical outcomes: An analysis of the Education Longitudinal Study of 2002 using a bioecological development model. *Bulletin of the Council for Research in Music Education*, *186*, 7–25.

Morrison, S. J. (1994). Music students and academic growth. *Music Educators Journal*, *81*(2), 33–36. doi:10.2307/3398812

National Association for Music Education. (2011a). *NAfME advocacy groundswell*. Retrieved from http://advocacy.nafme.org

National Association for Music Education. (2011b). *Students who study the arts score higher on SAT tests*. Retrieved from http://www.menc.org/supportmusic_cases/view/687

Newsweek. (2011). Best American high schools: How we compiled the list. *Newsweek*. Retrieved from http://www.thedailybeast.com/newsweek/2011/06/20/best-american-high-schools-how-we-compiled-the-list.html

Reimer, B. (1999). Facing the risks of the "Mozart effect." *Music Educators Journal*, *86*(1), 37–43. doi:10.2307/3399576

Reimer, B. (2005). The danger of music education advocacy. *International Journal of Music Education*, *23*(2), 139–142.

Rogers, W. H. (1993). *Regression standard errors in clustered samples. Stata Technical Bulletin*, 13, 19–23.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Southgate, D. E., & Roscigno, V. J. (2009). The impact of music on childhood and adolescent achievement. *Social Science Quarterly*, *90*(1), 4–21. doi:10.1111/j.1540-6237.2009.00598.x

Steiner, P. M., Cook, T. D., Shadish, W. R., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods*, *15*(3), 250–267. doi:10.1037/a0018719

Stewart, C. (1991). *Who takes music? Investigating access to high school music as a function of social and school factors* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor.

Vaughan, K. (2000). Music and mathematics: Modest support for the oft-claimed relationship. *Journal of Aesthetic Education, 34*(3/4), 149–166.

Vaughan, K., & Winner, E. (2000). SAT scores of students who study the arts: What we can and cannot conclude about the association. *Journal of Aesthetic Education, 34*(3/4), 77–90.

Winner, E., & Cooper, M. (2000). Mute those claims: No evidence (yet) for a causal link between arts study and academic achievement. *Journal of Aesthetic Education, 34*(3/4), 11–75.

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

Zwick, R., & Greif Green, J. (2007). New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. *Journal of Educational Measurement*, *44*(1), 23–45. doi:10.1111/j.1745-3984.2007.00025.x

Author Biography

Kenneth Elpus is assistant professor of music education at the University of Maryland. His research interests include music in education policy, music education as a context for adolescent development, and the analysis of large education data sets.

Submitted January 22, 2012; accepted October 19, 2012