education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal

Arizona State University

Volume 27 Number 138 October 28, 2019 ISSN 1068-2341

The Influence of School Accountability Incentives on Advanced Placement Access: Evidence from Pennsylvania

Paul Beach

Keith Zvoch
&
Michael Thier
University of Oregon
United States

Citation: Beach, P., Zvoch, K., & Thier, M. (2019). The influence of school accountability incentives on Advanced Placement access: Evidence from Pennsylvania. *Education Policy Analysis Archives, 27*(138). https://doi.org/10.14507/epaa.27.4602

Abstract: This study employed hierarchical piecewise growth modeling and two interrupted time series models to examine the effect of introducing an Advanced Placement (AP) school accountability incentive on AP access in Pennsylvania. Specifically, we examined whether adoption of an advanced course access accountability indicator was associated with an increase in AP course offerings initially and in the three years after the policy intervention. We also analyzed if the indicator differentially affected schools we hypothesized as sensitive or nonsensitive to the policy and examined demographic differences between those school groups. Pennsylvania's AP accountability incentive was associated with an initial increase in schools' AP course offerings, but the trajectory of change during the post-policy intervention period did not differ from the pre-policy baseline period. Also, the sizeable gap between schools with the most and fewest AP course offerings did not narrow across time. Instead, the gap widened. Our results suggest that adoption of AP school accountability incentives may not be a long-term solution to improving AP access for all schools or narrowing disparities in

Journal website: http://epaa.asu.edu/ojs/
Facebook: /EPAAA
Twitter: @epaa_aape

Manuscript received: 4/12/2019
Revisions received: 7/26/2019
Accepted: 7/27/2019

access between schools. We call for examinations in other states to determine if, and under what conditions, AP accountability incentives increase AP course offerings while narrowing access disparities.
Keywords: school accountability; Advanced Placement; educational equity; interrupted time series

La influencia de los incentivos de rendición de cuentas en las escuelas en el acceso a cursos de Advanced Placement: Evidencia de Pennsylvania

Resumen: Este estudio empleó modelos de crecimiento y dos modelos de series de tiempo interrumpido para examinar el efecto de introducir un incentivo de rendición de cuentas escolar de Advanced Placement (AP) en el acceso a AP en Pensilvania. Específicamente, examinamos si la adopción de un indicador avanzado de rendición de cuentas de acceso al curso se asoció con un aumento en la oferta de cursos AP inicialmente y en los tres años posteriores a la intervención de la política. También analizamos si el indicador afectó de manera diferencial a las escuelas que planteamos como sensibles o no sensibles a la política y examinamos las diferencias demográficas entre esos grupos escolares. El incentivo de responsabilidad AP de Pensilvania se asoció con un aumento inicial en la oferta de cursos AP de las escuelas pero la trayectoria del cambio desde el período anterior a la política y la intervención posterior a la política. Además, la brecha considerable entre las escuelas con la mayoría y la menor cantidad de ofertas de cursos AP se amplió con el tiempo. Nuestros resultados sugieren que la adopción de incentivos de rendición de cuentas escolar AP puede no ser una solución a largo plazo para mejorar el acceso AP para todas las escuelas o reducir las disparidades en el acceso entre las escuelas. Solicitamos exámenes en otros estados para determinar si y como estos incentivos aumentan las ofertas de cursos AP y reducen las disparidades de acceso.
Palabras-clave: rendición de cuentas; Advanced Placement; equidad educativa; series de tiempo interrumpidas

A influência dos incentivos de prestação de contas nas escolas no acesso à Advanced Placement: Evidências da Pensilvânia

Resumo: Este estudo empregou modelagem hierárquica de crescimento e modelos de séries temporais interrompidas para examinar o efeito da introdução de um incentivo de prestação de contas escolar de Advanced Placement (AP) no acesso ao AP na Pensilvânia. Especificamente, examinamos se a adoção de um indicador avançado de prestação de contas de acesso a cursos estava associada a um aumento nas ofertas de cursos de AP inicialmente e nos três anos após a intervenção política. Também analisamos se o indicador afetou diferencialmente as escolas que, segundo nossa hipótese, eram sensíveis ou não sensíveis à política e examinamos as diferenças demográficas entre esses grupos escolares. O incentivo de prestação de contas de AP da Pensilvânia foi associado a um aumento inicial nas ofertas de cursos de AP das escolas, mas a trajetória de mudança durante o período de intervenção pós-política não diferiu do período de linha de base pré-política. Além disso, a diferença considerável entre as escolas com o menor número de ofertas de cursos de AP aumentou. Nossos resultados sugerem que a adoção de incentivos à prestação de contas nas escolas de AP pode não ser uma solução de longo prazo para melhorar o acesso ao PA em todas as escolas ou diminuir as disparidades no acesso entre as escolas. Solicitamos exames em outros estados para determinar se e como esses incentivos aumentam as ofertas de cursos da AP e reduzem as disparidades de acesso.

Palavras-chave: prestação de contas; Advanced Placement; equidade educacional; séries temporais interrompidas

The Influence of School Accountability Incentives on Advanced Placement Access: Evidence from Pennsylvania

The Advanced Placement (AP) program, administered by the College Board, has garnered significant federal resources and considerable attention in state policy. For example, the U.S. Department of Education distributed $273 million in grants via the AP Incentive Program from 2002 to 2011 to increase AP access for students from socioeconomically disadvantaged backgrounds (U.S. Department of Education, 2013). The Federal government is not alone in its attempts to improve AP access and performance. According to the Education Commission of the States (2016), every state except Alaska, Kansas, Nebraska, and Vermont employed at least one AP policy, including 32 states that used AP participation or performance as high school accountability indicators. The purpose of this study is to examine the effects of school accountability incentives on AP access in Pennsylvania, the only state to our knowledge that has employed an accountability indicator designed to increase school-level AP course offerings.

The wide-scale use of AP indicators in school accountability stems in large part from the belief that AP accountability incentives will cause schools to provide greater access to the range of benefits associated with AP participation (Achieve & Jobs for the Future, 2015; The Education Trust, 2019), including the program's potential for providing students with academic, economic, and social advantages. When academic outcomes have been investigated, researchers affiliated with the College Board, as well as those working independently, have "generally found that AP participation is associated with higher academic achievement" (Warne, 2017, p. 4). The College Board (2014) attributes the academic benefits to the college-level knowledge and skills proffered through participation in one or more of its 38 courses and associated exams.[1] AP participation also positions students to receive economic benefits, such as gaining college credit for scoring high enough on AP exams, reducing time and expense en route to college graduation (Ackerman, Kanfer, & Calderwood, 2013). A less discussed but equally important advantage of AP is the mark of distinction conferred on students and schools (Klugman, 2013). Students, especially those seeking admission to selective universities, can leverage their AP participation as an indicator of college readiness on their high school transcripts (Geiser & Santelices, 2004; Sadler, 2010). Relatedly, many school leaders seek to bolster their AP course offerings and exam performance to gain a competitive edge over other schools (Klugman, 2013). Competition among schools is facilitated partly by ranking systems such as the *U.S. News and World Report Best High School Rankings*, where AP participation and performance indicators account for 40% of the composite score for more than 17,000 high schools (Morse & Brooks, 2019).

Advocates for the inclusion of AP indicators in school accountability also argue access to advanced courses that include the possibility of earning college credit "matters to students and parents" (Achieve & Jobs for the Future, 2015, p. 2; The Education Trust, 2019, p. 2). The steady increase in school-level AP access and exam participation across time supports the notion that parents and students demand greater AP access. Between academic years 1996-97 and 2015-16, the

[1] P*articipation* means enrolling and completing an Advanced Placement (AP) course. *Performance* means students' scores on AP exams.

percentage of U.S. public high schools[2] administering at least one AP exam increased from 55% to 73% (College Board, 1997; 2016). In 2016-17 alone, more than 2.3 million—or 16% of U.S. public high school students[3]—sat for 4.2 million AP exams (College Board, 2017), though the number of students participating in AP is likely higher because some students elect not to sit for AP's end-of-course exams. AP access, however, is not equal across schools. AP access is conditioned on several school-level characteristics, including socioeconomic status, racial/ethnic composition, enrollment, student-teacher ratio, and geographic locale (Thier, Beach, Todd, & Coleman, 2016). The few studies to examine growth in AP access have found schools that serve higher proportions of students from socioeconomically disadvantaged backgrounds and/or students of color increase AP course offerings at slower rates than schools with student bodies of higher socioeconomic status and/or less diversity (Klopfenstein, 2004; Klugman, 2013; Zarate & Pachon, 2006). These inequalities are especially problematic considering the AP program has been repositioned to simultaneously improve college access and better prepare students for postsecondary success.

In light of ongoing federal and state investments, continued growth in AP participation, and persistent equity challenges, this study examined if Pennsylvania's accountability reform led to improved AP access. Pennsylvania reformed its school accountability system in 2013 after having its Elementary and Secondary Education Act flexibility waiver approved by the U.S. Department of Education. Pennsylvania's new accountability system included a range of new indicators, including one that awarded schools points based on the number of advanced courses offered, such as AP. To our knowledge, Pennsylvania is the only state that has employed an accountability indicator specifically designed to improve AP access as states typically adopt AP indicators that focus on exam performance (Conley, Thier, Beach, Lench, & Chadwick, 2014). Additionally, the design of Pennsylvania's AP access indicator appeared to target schools with the fewest advanced course offerings. This study was thus designed to ascertain if including an AP access indicator in Pennsylvania's school accountability system was associated with an increase in the number of AP courses offered across schools initially and during a three-year policy intervention period, with a specific focus on schools with the fewest course offerings.

Previous Research

Several authors have examined AP access, including what student and school demographic variables predict access as well as what within-school barriers prevent access for students (e.g., Conger, Long, & Iatarola, 2009; Klopfenstein, 2004; Thier et al., 2016). Although these authors and others have put forth different theories for explaining why disparities in AP access exist between and within schools, it is indisputable that certain schools and students are more likely than others to have exposure to the potential benefits AP can provide (Kolluri, 2018). State policymakers have instituted numerous AP-specific policies to address these disparities, and/or in some cases, state AP accountability policy has been enacted to simply expand opportunities for students to access AP's potential benefits (Education Commission of the States, 2006; 2016). Yet, only a handful of studies have specifically examined the effects of state AP policies. The current study adds to the literature by

[2] We obtained school-level codes from the National Center for Education Statistics' Elementary/Secondary Common Core of Data, which we used to determine the number of public high schools in academic years 1996-97 through 2015-16 that administered at least one AP exam. We excluded schools coded as *other* because some of these schools may not offer the high school grades in which students typically enroll in AP courses.

[3] The percentage of public school students that sat for at least one AP exam in 2016-17 is based on the number of public high school students in 2015-16, which was approximately 14 million according to the National Center for Education Statistics' Elementary/Secondary Common Core of Data.

using quasi-experimental methods to quantitatively investigate the effects of a state-level AP policy change. We begin with a review of past research on AP at the school level, as well as studies examining policy interventions designed to improve AP access and performance. We then describe how economic theory can help explain the design and intended effect of school accountability incentives before describing Pennsylvania's recent reforms.

School-Level Advanced Placement Access

AP originated in the 1950s as a program designed to provide elite high school students with access to college-level curriculum (Lacy, 2010). The AP program was accessed primarily by high-achieving students until the 1980s and 1990s when Schneider (2009) argues the general popularity of the program, state policy support, and several initiatives that provided subsidies for AP exam fees led to substantial increases in access for all students. AP access continued to expand for all students during the next two decades. For example, data from two nationally representative samples shows the percentage of students who received AP credit on their high school transcripts more than doubled for Asian, Black, Hispanic, and White students from 1990 to 2013 (Malkus, 2015). However, the gap in AP participation between these groups remained steady across time, with Asian students (70% in 2013) the most likely to earn AP credit, followed by White (41%), Hispanic (37%), and Black (27%) students. The College Board (2018) has addressed these equity concerns by encouraging the expansion of AP access through the "elimination of barriers that restrict access to AP for students from ethnic, racial, and socioeconomic groups that have been traditionally underrepresented" (p. 9). However, disparities in access based on race/ethnicity and socioeconomic status persist despite the steady growth in AP access for all students (Kolluri, 2018).

Several researchers have found AP access to associate negatively with schools' percentage of students eligible for free or reduced price lunch (Barnard-Brak, McGaha-Garnett, & Burley, 2011; Iatarola, Conger, & Long, 2011; Klopfenstein, 2004; Zarate & Pachon, 2006). There also seems to be some agreement across studies that schools with a high proportion of White students have more AP access (Barnard-Brak et al., 2011; Cisneros, Gomez, Powers, Holloway-Libell, & Corley, 2014; Iatarola et al., 2011) than schools with a high proportion of Black, Latinx, or American Indian/Alaska Native students (Conger et al., 2009; Darity, Castellino, Tyson, Cobb, & McMillen, 2001; Klugman, 2013), although authors vary in how they categorize student groups. With respect to school size, the number of students enrolled has a positive relationship with AP access (Cisneros et al., 2014; Iatarola et al., 2011; Jeong, 2009; Malkus, 2016; Zarate & Pachon, 2006). The few studies that have examined relations between AP access and schools' student-teacher ratios produced negative (Iatarola et al., 2011) or inconclusive associations (Conger et al., 2009). Finally, regarding geographic locale, schools in suburban areas, on average, appear to have more AP access than schools in cities, towns, or rural areas (Gagnon & Mattingly, 2016; Jeong, 2009; Klopfenstein, 2004; Klugman, 2013; Zarate & Pachon, 2006).

Advanced Placement Policy Analyses

A relatively small group of researchers have examined state policies designed to improve AP access and performance. For example, Jeong (2009) used the Education Longitudinal Study of 2002 to analyze two types of state policies: AP exam fee waivers and merit-based incentives (e.g., scholarships for students, cash for teachers) tied to AP exam scores. Jeong found students in the nine states that waived exam fees were more likely to take AP exams than peers in other states, but students in the five states with merit-based policies were not more likely to take AP exams. A separate study in Texas found that subsidies for AP exam fees "provided no incentive for schools to expand their AP course offerings" and that high-poverty and rural schools continued to offer far less AP access than other schools (Klopfenstein, 2004, p. 10).

Jackson (2010) analyzed how the Texas AP Incentive Program affected five outcomes, including AP course participation and performance, SAT/ACT participation and performance, and college matriculation. Texas' program provided cash incentives to teachers and students for passing AP exam scores (i.e., typically 3 or higher on a 1-5 scale). Using a differences-in-differences approach, Jackson compared schools that would eventually adopt the program and those that never did, attributing to the AP Incentive Program improvements in all outcomes but SAT/ACT performance. Similarly, Kramer (2016) analyzed whether state merit-based incentives led to increases in AP course participation and exam-taking. During calendar years 1993-2005, 14 states instituted policies that paid for all or some of a student's college tuition based on high school grade-point average. States where grade-point average included weighted bonuses for AP course-taking saw higher increases in course and exam participation than states that did not add weight for AP courses.

Of the studies examining the effects of AP policy, Klugman (2013) was alone in finding that school-level inequalities in AP access persisted, and in some cases accelerated, despite a "short lived" intensive investment aimed in part at equalizing AP access across schools (p. 4). Klugman explored whether three California policies designed to improve AP access narrowed inequalities associated with school-level socioeconomic status and racial/ethnic composition. Those policies included AP Challenge Grants (e.g., funding for teacher training, instructional materials, and tutoring services), funding for AP-specific training and support in the Advancement Via Individual Determination program, and a University of California initiative aimed at increasing online AP course offerings. Klugman used 3-level hierarchical linear models to analyze 10 years of data, finding evidence that California's intervention increased AP course offerings and course participation, but "did very little to decrease inequalities" (p. 26). Schools predominately serving students who were Asian or White saw larger increases in AP course offerings and participation than those predominately serving students who were Black or Latinx.

Using a resource deprivation framework and the theory of Effectively Maintained Inequalities, Klugman (2013) explained why California's multi-pronged policy approach failed to close gaps in AP access. First, Klugman argues that scholars' typical explanations when examining school-level AP access inequalities focus on the resource constraints facing families and schools that serve high percentages of students traditionally marginalized in public education based on socioeconomic status and race/ethnicity. Resource constraints can be material (e.g., teachers with advanced degrees) or immaterial (e.g., social and cultural capital) in nature (Klugman, 2013; Iatarola et al., 2011). Under the resource deprivation framework, interventions targeting constraints that schools, families, and students face should help reduce inequalities in school-level AP access, thus improving opportunities to learn advanced coursework.

The theory of Effectively Maintained Inequalities (Lucas, 2001) posits that students from high socioeconomic families, who are increasingly concentrated in high socioeconomic school contexts (Owens, Reardon, & Jencks, 2016), will demand more AP courses to gain a competitive edge in admissions offices at elite colleges and universities. Klugman (2013) follows by arguing that schools will continually increase AP course offerings to match demands from high socioeconomic families to produce more opportunities for their children as they pursue marks of distinction for college admissions. Despite California's effort to provide resources to improve access for schools with the fewest AP course offerings, schools with already high levels of AP access kept pace or increased course offerings relative to other schools, thus maintaining advantages.

As the Education Commission of the States (2016) shows, states also employ AP policies that do not provide direct resources to schools or students, such as school accountability incentives. For example, Rowland and Shircliffe (2016) examined one school's reaction to the Florida Partnership for Minority and Underrepresented Student Achievement Act of 2007 and the state's decision to use AP participation and exam performance data to rank schools in its accountability

system in 2009. The resource-focused Florida Partnership for Minority and Underrepresented Student Achievement Act provided professional development to teachers in the core content areas and assisted school districts on identifying minority and underrepresented students for potential participation in AP. Interviews with teachers, district leaders, and elected officials revealed tensions between simultaneous efforts to expand AP access as part of the Florida Partnership for Minority and Underrepresented Student Achievement Act and improve AP exam performance in the state's school accountability system. The authors found that Florida's accountability system complicated "teachers' views of AP expansion, particularly in relation to the use of pass rates to evaluate AP instruction" and in the calculation of school accountability ratings (p. 415). Because expanding AP access might yield lower AP exam scores—absent educational inputs that concurrently focus on equity and excellence—some states' approaches to school accountability can disincentivize school staff from recruiting more students for AP courses. The same approach of restricting AP access to only high performers would reward schools with better accountability ratings. Below, we describe how economic theory can help explain why schools, such as the one in Rowland and Shircliffe's study, may be sensitive to school accountability incentives that aim to improve AP access.

Principal Agent Theory and School Accountability

Economists developed principal-agent theory to help explain contractual relations between managers and employees in the business world (Holmstrom & Costa, 1986; Jensen & Meckling, 1976; Milgrom & Roberts, 1988). Not long after it was conceived, scholars began modifying principal-agent theory to better conceptualize relational and contractual dynamics between different actors in the public sector (Moe, 1984), including education (e.g., Bae, 2018; Figlio & Loeb, 2011; Moe, 2003; Polikoff, McEachin, Wrabel, & Duque, 2014). As Bae (2018) describes, in the basic model, "educators (agents) might not act in accordance with the interests of stakeholders (e.g., parents, community members, and policymakers – the principal) unless motivated to do so with incentives (both positive and negative) created through accountability systems" (p. 4). Scholars vary in who they think are sensitive to these incentives, but at the most basic level it appears school accountability policy is designed to motivate teachers and administrators to improve their educational practices in accordance with the goals set by policymakers (Figlio & Loeb, 2011). Moreover, "school-based accountability operates on the notion that incentives and public pressure from publicly reported information will result in improved student outcomes" (Bae, 2018, p. 5).

Under the No Child Left Behind Act of 2001, federal policymakers set the goal of universal proficiency in numeracy and literacy and required states to publicly report school-level standardized test scores in mathematics and reading as a high-stakes accountability incentive. No Child Left Behind was heavily critiqued for focusing too narrowly on numeracy and literacy proficiency at the expense of other outcomes that stakeholders also value, such as civic engagement, the Arts, and social skills (e.g., Rothstein, Jacobsen, & Wilder, 2008). Part of this critique can be explained by principal-agent theory, which predicts policies will have unintended consequences when agents (i.e., educators) are required to multi-task on the job and pursue multiple goals, but are only held accountable by principals (i.e., policymakers) to a narrow set of outcomes (Holmstrom & Milgrom, 1991). In these cases, agents will react by focusing a disproportionate amount of energy on improving the outcome(s) they are held accountable to while paying less attention to other priorities. Indeed, research shows schools largely responded to high-stakes school accountability policy by focusing disproportionately on the tested subjects used for accountability purposes, also referred to as curriculum narrowing (Berliner, 2011).

Perhaps the most common policy recommendation stemming from the unintended consequence of curriculum narrowing was to expand the type and number of educational indicators used to hold schools accountable. For example, Darling-Hammond, Wilhoit, and Pittenger (2014)

argue ideal state accountability systems should provide adequate resources to schools, invest in building the capacity of teachers and administrators, and create a data dashboard that includes multiple indicators of school performance to "provide a better accounting of what schools are doing and with what results" (p. 27). Darling-Hammond et al. go on to argue these types of multiple indicator systems existed in several states during the 1990s, went relatively dormant during the 2000s due to No Child Left Behind requirements, and eventually made a comeback through the Elementary and Secondary Education Act (ESEA) flexibility waiver process.

ESEA waivers allowed state policymakers the opportunity to request flexibility on certain No Child Left Behind requirements, including the indicators used to hold schools accountable. Several states seized the opportunity provided by ESEA waivers to expand the type and number of indicators used to assign school quality ratings and identify schools in need of improvement. For example, Polikoff et al. (2014) analyzed 42 approved ESEA waivers and found that 38 states created some sort of composite index that employed multiple indicators to assign school quality ratings. Pennsylvania, one of these 38 states, used the ESEA waiver process to create a composite index that employed 23 indicators to generate school quality ratings for high schools.

Pennsylvania's Accountability System Redesign

Pennsylvania policymakers first presented plans to redesign the state's school accountability system in February 2013 upon submitting its initial ESEA flexibility waiver to the U.S. Department of Education. Under its flexibility waiver, Pennsylvania created the *School Performance Profile*. For high schools, the *School Performance Profile* included 23 indicators across five categories: academic achievement, closing achievement gaps, academic growth, other academic indicators, and extra-credit indicators. Approved state flexibility waivers, including Pennsylvania's, expired after 2017-18 when the Every Student Succeeds Act took effect. As a result, Pennsylvania schools have five years (2012-13 to 2017-18) of "Building Level Academic Score" data in the *School Performance Profile*.

The Building Level Academic Scores for high schools relied on data from 18 of the 23 accountability indicators to generate a score on a 100-point scale (schools could earn extra credit accountability points from the five remaining indicators). One of these 18 indicators measured schools' college credit-bearing course offerings (e.g., AP, International Baccalaureate) in four core academic areas: mathematics, English language arts, life and physical sciences, social studies and history. This indicator, which we refer to as "Access to Advanced Coursework" (AAC), accounted for 2.50% of schools' Building Level Academic Scores. This study specifically analyzes the effect of introducing the AAC indicator into Pennsylvania's accountability system on school-level AP access.

An important feature of the AAC indicator was that schools earned maximum accountability points when they had at least one student enroll in an advanced course in each of the four core academic areas (e.g., AP Statistics, AP English Literature and Composition, AP U.S. History, and a dual-enrollment course in chemistry). Once schools reached the maximum AAC points there was no accountability point incentive for offering more advanced courses. We refer to these schools as *policy nonsensitive* because they had already earned maximum accountability points when the new accountability system took effect, and the other schools as *policy sensitive* because they had not yet received maximum points. Stated differently, policy sensitive schools had more of an accountability incentive to increase AP course offerings because they could earn more accountability points from doing so whereas policy nonsensitive schools could not.

Principal-agent theory predicts publicly reporting school performance data will motivate educators to improve on the outcomes being measured. Thus, we designed our study to examine if introduction of Pennsylvania's AP accountability incentive increased the average number of AP courses offered initially and across time. Of particular interest was ascertaining whether changes in Pennsylvania's school accountability system led to relatively more AP course offerings in policy

sensitive schools, thereby narrowing AP access disparities between policy sensitive and nonsensitive schools. Note, however, the primary intent of the current investigation was to determine if inclusion of the AAC indicator had the intended effect of increasing AP courses offerings. As such, we employed school demographics as statistical controls in our analytical models to ensure any changes we observed in AP course offerings was not confounded by changes in school characteristics.

Method

The present study examined whether the introduction of the AAC indicator in Pennsylvania's school accountability system was associated with changes in school-level AP course offerings. We employed hierarchical piecewise growth modeling and two interrupted time series (ITS) models, simple and comparative (CITS), to examine whether the AAC indicator led to an average increase in AP access initially and during a three year policy intervention period (Raudenbush & Bryk, 2002; Shadish, Cook, & Campbell, 2002; Singer & Willet, 2003). We also analyzed if policy sensitive and nonsensitive schools had different trajectories in AP course offerings prior to adoption of the AAC indicator, if adoption of the AAC indicator differentially affected the two groups, and if the two groups differed with respect to five school demographic characteristics.

Data

We utilized three data sources in this study: the Pennsylvania *School Performance Profile* website data files, the National Center for Education Statistics' Elementary/Secondary Common Core of Data, and the AP Course Audit database. In the Pennsylvania *School Performance Profile*, 638 schools have data for the AAC indicator for five years: one year before the accountability policy change (2012-13) and the four years after (i.e., until 2016-17). We matched schools from the Pennsylvania *School Performance Profile* with those in the Common Core of Data based on school names, cities, and addresses, failing to match only one school. We excluded 82 schools in the Common Core of Data coded as vocational (n = 18), alterative (n = 1), or other (i.e., not high schools; n = 63). For each school under analysis, we obtained the following variables from the Common Core of Data: proportion of students eligible for free or reduced priced lunch (variable label, *poverty*), proportion of students who are White (*White*), total number of enrolled students (*enroll*), student-teacher ratio (*ratio*), and geographical locale (categorized as *city, suburb, fringe, distant,* and *remote*). Finally, we merged our working version of those datasets with the AP Course Audit database using school names, cities, and addresses. The AP Course Audit database, administered by Inflexion (formerly the Educational Policy Improvement Center), includes the number and type of AP courses offered in all schools starting from 2007-08. In all, we retained an analytical sample of n = 555 high schools that received AAC ratings in Pennsylvania's accountability system. Our analytical sample includes 99.82% of the traditional high schools theoretically affected by Pennsylvania's AAC indicator.

Research Design

We used 2-level hierarchical piecewise growth models with time nested in schools to analyze the AAC indicator's effect on school-level AP course offerings initially (i.e., changes in intercept) and during a three year policy intervention period (i.e., changes in slope). Hierarchical piecewise growth modeling allowed us to estimate two growth functions, one for the pre-policy intervention period (2007-08 to 2013-14) and one for the post-policy intervention period (2014-15 to 2016-17). We did not develop a 3-level model because of limited variability within districts: 94.67% of the districts in our sample contained a single high school.

As discussed briefly above, the design of the AAC indicator created two distinct groups of schools, policy sensitive and nonsensitive. The AAC indicator operated on a five-level scale (0, 25,

50, 75, 100), and schools could earn no additional accountability points once they reached the mathematical maximum of 100 AAC points. Policy sensitive schools did not earn all 100 points on the AAC indicator in 2012-13, the second to last pre-policy intervention year. Policy nonsensitive schools earned all 100 AAC points in 2012-13, so they could not earn more accountability points by subsequently increasing AP course offerings. The comparative aspect of the time series design examined whether policy sensitive and nonsensitive schools exhibited a differential response after Pennsylvania introduced the AAC accountability indicator.

Measures

In Table 1, we present descriptive statistics for all the variables in our study. We left all variables uncentered as each had a true zero. For continuous variables, we report means, standard deviations, ranges, and skewness, along with proportions for all five geographic local subcategories. The single missing data point in the dataset came from the student-teacher ratio variable; we dropped this school from the independent sample *t*-tests that analyzed whether policy sensitive and nonsensitive schools differed with respect to student-teacher ratio.

Table 1
Descriptive Statistics for Dependent and School Demographic Variables (n = 555)

Continuous	*M*	*SD*	Range	Skew
AP courses[a]	7.78	6.73	0-29	0.82
Poverty[b]	0.39	0.22	0-1	0.98
White[b]	0.78	0.28	0-1	-1.70
Enrollment[b]	868.21	558.78	92-3638	1.54
Ratio[b]	14.68	2.96	7-37	2.09
Categorical	%			
City[b]	14.6			
Suburb[b]	41.1			
Fringe[b]	20.5			
Distant[b]	19.6			
Remote[b]	4.1			

Note. [a] Dependent variable: data aggregated from academic years 2007-08 to 2016-17; [b] School demographic variables: data only from academic year 2013-14; Poverty = a school's aggregated percentage of students eligible for free or reduced priced lunch; White = White students' proportion of a school's enrollment; Enrollment = the number of students reported by a school; Ratio = a school's enrollment divided by its number of full-time equivalent classroom teachers.

Dependent, growth, intervention, and treatment group variables. Our dependent variable for all analyses was the number of AP courses (e.g., AP Biology, AP Spanish) the College Board authorized a school to offer in a given academic year. The range for *AP courses* across the high schools in our sample was from 0 to 29 courses. We chose to use AP course offerings as our dependent variable instead of AAC points for two reasons. First, AP course offerings allowed us to examine the complete range of AP access across all schools. The AAC indicator has an artificial ceiling, making it difficult to determine how advanced course access changed in policy nonsensitive schools before or after the policy intervention. Second, we used 10 years of AP Course Audit data to estimate two AP access growth functions, one prior to and one after the policy intervention. Conversely, the lack of pre-intervention AAC indicator data made it impossible to calculate pre-policy intervention growth trajectories.

We created time, intervention, and treatment group variables for the hierarchical piecewise growth models. We analyzed growth in school-level AP access using a linear term, referred to as *Pre TxSlope*, signifying the pre-policy slope. We coded Pre TxSlope as 0 in academic year 2007-08 (i.e., the intercept), 1 for 2008-09, 2 for 2009-10, and so on. We employed the variable *TxLevel* to examine intercept changes in AP courses at the point of policy intervention and *Post TxSlope* to examine slope changes in AP courses during the post-policy intervention timeframe. TxLevel was coded as 0 prior to the policy intervention (academic year 2013-14), and 1 afterward. Post TxSlope was coded as 0 through the first year of the policy intervention (2014-15), 1 for the second year (2015-16), and 2 for the third year (2016-17). We considered 2014-15 the first policy intervention year because Pennsylvania released the first round of school quality ratings in the beginning of the 2013-14 academic year using data from 2012-13. We chose not to use 2013-14 as the first policy intervention year because as Klugman (2013) argues, district and school administrators make decisions on what AP courses to offer the year prior to when they are offered. Finally, *TxGroup* is a dichotomous variable used to classify schools into policy sensitive and nonsensitive groups.

School demographic variables. We used school-level variables for two purposes. First, we used four continuous and one categorical variable to examine demographic differences between policy sensitive and nonsensitive schools. We treated all variables in this first analysis as time-static using data from 2013-14, the last pre-policy intervention year. This approach helps contextualize the general demographic makeup of policy sensitive and nonsensitive schools in the academic year before the policy intervention was enacted. Second, we added the four continuous variables as time-varying covariates in our hierarchical piecewise growth model.

The four continuous variables used in both analyses include: poverty, White, enroll, and ratio. Poverty is the aggregated percentage of students within a school eligible for free or reduced priced lunch. Schools' racial/ethnic composition was expressed by computing the number of White students as a proportion of school enrollment. Enrollment is simply the total number of students reported by each school. Our last continuous variable was student-teacher ratio: enrollment divided by the number of full-time equivalent classroom teachers. Poverty and racial/ethnic composition are frequently used variables in examinations of educational equity. Enrollment and student-teacher ratio are useful variables because they approximate schools' contextual needs for breadth of AP course offerings as well as schools' resources to meet such needs (see Thier et al., 2016).

We used the lone categorical variable, geographic locale, to examine differences in where policy sensitive and nonsensitive schools are situated relative to urban areas. To incorporate geographical locale, we modified the National Center for Education Statistics' Urban-Centric Locale Codes to create five dummy-coded variables: city, suburb, fringe, distant, and remote. The Urban-centric codes feature four broad categories: city, suburb, town, and rural, but this approach masks important nuances between distant, fringe, and remote subcategories of towns and rural areas (Thier, Beach, Martinez, & Hollenbeck, 2017). Analyses that include only city, suburb, and rural designations might also fail to detect certain locales that differentiate sites of advantage from sites of disadvantage. Therefore, we combined the three subcategories each for city and suburb. Then, we recategorized the three town or rural sublevels into fringe (combining town and rural fringe locales), distant (town and rural distant locales), and remote (town and rural remote locales)—to prioritize urban proximity over the quartiles' somewhat arbitrary community population counts.

Analyses

First, we used independent sample *t*-tests and chi-square analyses to examine demographic differences between policy sensitive and nonsensitive schools. We used independent sample *t*-tests to determine if there were statistical differences between policy sensitive and nonsensitive schools

on the dependent variable (AP courses) and the four continuous school characteristics (poverty, White, enroll, and ratio). Due to the nature of the geographic local variable, we used chi-square analyses to examine differences between policy sensitive and nonsensitive schools regarding the five geographical locale categories (city, suburb, fringe, distant, and remote). For both sets of analyses, we set $\alpha = .05$ and used Bonferroni's procedure to control for familywise Type I error.

Second, we examined whether the introduction of Pennsylvania's AAC indicator resulted in statistically significant changes in school-level AP course offerings initially (i.e., change in intercept from pre- to post-policy intervention year) and during the three years of policy intervention (i.e., change in slope from pre- to post-policy intervention years) through application of hierarchical piecewise growth modeling. Our analyses also allowed us to determine how policy sensitive and nonsensitive schools differed with respect to the initial level of AP course offerings (i.e., intercept) and changes across time (i.e., pre-intervention growth trajectory). We employed a four-step model-building process to determine whether Pennsylvania's AAC indicator led to statistically significant increases in AP course offerings on average for all schools and differentially for policy sensitive schools. The first and second steps in our model-building process evaluated whether sufficient between-school variation existed to warrant use of hierarchical linear modeling and whether growth in school-level AP course offerings should be modeled linearly. Our model-building process confirmed that hierarchical linear modeling was appropriate for our dataset (i.e., intraclass correlation coefficient of 0.90) and that AP course offerings should be modeled linearly (i.e., quadratic and cubic growth terms were not statistically significant, $p > .05$).

Next, the simple ITS model (see Equation 1) included the outcome variable, the number of AP courses offered, and terms used to estimate the piecewise growth trajectory: Pre TxSlope, TxLevel, and Post TxSlope. The simple ITS model provided initial evidence of whether Pennsylvania's AAC indicator affected average AP course offerings across all schools.

$$
\begin{aligned}
(AP\ Courses)_{ti} &= \pi_{0i} + \pi_{1i}(Pre\ TxSlope)_{ti} + \pi_{2i}(TxLevel)_{ti} + \pi_{3i}(Post\ TxSlope)_{ti} + e_{ti} \\
\pi_{0i} &= \beta_{00} + r_{0i} \\
\pi_{1i} &= \beta_{10} + r_{1i} \\
\pi_{2i} &= \beta_{20} + r_{2i} \\
\pi_{3i} &= \beta_{30} + r_{3i}
\end{aligned} \tag{1}
$$

In the CITS model (see Equation 2), we estimated cross-level interactions by adding the dummy-coded treatment group variable (policy sensitive and nonsensitive) to the model. The CITS model allowed us to examine if schools reacted differentially to adoption of the AAC indicator based on their group classification.

$$
\begin{aligned}
(AP\ Courses)_{ti} &= \pi_{0i} + \pi_{1i}(Pre\ TxSlope)_{ti} + \pi_{2i}(TxLevel)_{ti} + \pi_{3i}(Post\ TxSlope)_{ti} + e_{ti} \\
\pi_{0i} &= \beta_{00} + \beta_{01}(TxGroup) + r_{0i} \\
\pi_{1i} &= \beta_{10} + \beta_{11}(TxGroup) + r_{1i} \\
\pi_{2i} &= \beta_{20} + \beta_{21}(TxGroup) + r_{2i} \\
\pi_{3i} &= \beta_{30} + \beta_{31}(TxGroup) + r_{3i}
\end{aligned} \tag{2}
$$

Finally, we estimated a covariate-adjusted CITS model to control for changes in school demographics across time. Controlling for school demographics allowed us to probe the alternative theory that shifts in school demographics—not the introduction of the AAC indicator—could explain any observed changes in AP course offerings. To probe this internal validity threat, we added four time-varying covariates to Level 1 in the piecewise hierarchical growth model: poverty,

racial/ethnic composition, enrollment, and student-teacher ratio. We excluded geographic locale since this variable is stable across time and therefore could not be responsible for changes in AP course offerings from the pre- to post-policy intervention period. We used R's lme4 package to analyze all three models (Bates, Maechler, Bolker, & Walker, 2014).

Results

We begin this section by reporting the statistically significant differences between policy sensitive and nonsensitive schools found on most demographic variables. Next, we present hierarchical piecewise growth modeling analyses regarding the effects of Pennsylvania's AAC indicator. Our ITS analyses determined whether the introduction of the AAC indicator was associated with statistically significant changes in the intercept and post-policy slope. We found Pennsylvania's school accountability reforms resulted in an initial—but not sustained—increase in AP course offerings across all schools. Additionally, policy sensitive and nonsensitive schools did not appear to react differentially to the introduction of the AAC indicator: both school types initially increased AP course offerings and then returned to their respective pre-policy intervention trajectories. Importantly, these results showed that sizeable gaps between policy sensitive and nonsensitive schools—schools with the most and fewest AP course offerings—did not narrow, and instead appeared to widen across time. Finally, we found that controlling for changes in school demographics did not alter the direction or statistical significance of the policy intervention effect.

Demographic Differences between Policy Sensitive and Nonsensitive Schools

In Table 2, we present the demographic differences between policy sensitive and nonsensitive schools. In 2013-14—the last pre-policy intervention year—the 162 policy sensitive schools offered an average of 2.42 AP courses, 7.81 fewer than the average among the 393 policy nonsensitive schools. On average, policy nonsensitive schools enrolled fewer students from socioeconomically disadvantaged backgrounds and more White students. Enrollment among policy nonsensitive schools was nearly double that of policy sensitive schools. Group differences regarding mean student-teacher ratios were not statistically significant. In terms of geographic locale, policy sensitive schools had more schools located in cities as well as in distant and remote locales. By contrast, policy nonsensitive schools had more schools in suburbs. Finally, both groups had nearly the same proportion of schools located in fringe areas.

Table 2
Empirical Contrasts of Policy Sensitive (n = 162) and Nonsensitive (n = 393) Schools

t tests (*df*)	Treatment group	*M*	*SD*	*t*	95% CI
AP courses (553)	Sensitive	2.42	3.01	14.30*	6.74, 8.88
	Nonsensitive	10.23	6.67		
Poverty (553)	Sensitive	0.52	0.23	-9.27*	-0.21, -0.14
	Nonsensitive	0.34	0.19		
White (553)	Sensitive	0.70	0.38	4.60*	0.07, 0.17
	Nonsensitive	0.82	0.22		
Enroll (553)	Sensitive	515.16	268.08	10.45*	404.85, 592.32
	Nonsensitive	1013.74	582.25		
Ratio (552)	Sensitive	14.63	4.00	0.25	-0.48, 0.61
	Nonsensitive	14.70	2.41		

Table 2 cont.
Empirical Contrasts of Policy Sensitive (n = 162) and Nonsensitive (n = 393) Schools

χ^2 tests (*df*)	Treatment group	Proportion	χ^2
City (1)	Sensitive	0.27	26.20*
	Nonsensitive	0.10	
Suburb (1)	Sensitive	0.17	53.53*
	Nonsensitive	0.51	
Fringe (1)	Sensitive	0.21	0.03
	Nonsensitive	0.20	
Distant (1)	Sensitive	0.27	6.91*
	Nonsensitive	0.17	
Remote (1)	Sensitive	0.09	11.65*
	Nonsensitive	0.02	

Note. *$p < .05$; Bonferroni's correction accounts for Type I error; Data are from academic year 2013-14; *df* = degrees of freedom; *M* = mean, *SD* = standard deviation; *t* = t-statistic; *CI* = confidence interval

Policy Intervention and Treatment Effects

In Table 3, we present findings for the simple ITS, CITS, and covariate-adjusted CITS models. In the simple ITS, both Pre TxSlope (β_{10}) and TxLevel (β_{20}) terms were statistically significant, $p < .05$. On average, school-level AP course offerings (Pre TxSlope) increased by 0.21 courses per year prior to the AP policy intervention. The TxLevel finding suggests that Pennsylvania's AAC policy intervention led to an initial average change of 0.34 AP course offerings per school. However, the Post TxSlope intervention variable (β_{30}) was not statistically significant, $p > .05$. In sum, aside from the immediate increase in the number of AP courses, growth trajectories in AP course offerings did not change from the pre- to the post-policy intervention period, on average.

In the CITS model, the estimate associated with the treatment group variable (β_{01}) shows that during the first year of the study period (i.e., academic year 2007-08), policy sensitive schools offered 6.67 fewer AP courses than policy nonsensitive schools, $p < .05$. In addition, the interaction between the treatment group variable and the linear growth term (β_{11}) revealed the slope for policy sensitive schools was statistically different than for policy nonsensitive schools across the baseline period, $p < .05$. The slope for TxGroup suggests policy nonsensitive schools increased AP course offerings at a rate of 0.20 more courses per year than policy sensitive schools prior to the AP policy change. That is, the gap between policy sensitive and nonsensitive schools widened from academic years 2007-08 to 2014-15. However, neither interaction between the treatment group and the two post-policy time function variables (β_{21} and β_{31}) resulted in statistically significant findings, $p > .05$. In other words, there was no statistical difference in how policy sensitive and nonsensitive schools reacted to the introduction of the AAC indicator. Both school groups initially increased AP course offerings after introduction of the AAC indicator and then returned to their respective pre-policy intervention growth trajectories.

Table 3
Results for the Simple ITS, CITS, and Covariate-Adjusted CITS Hierarchical Piecewise Growth Models

	Simple ITS	CITS	Covariate-Adjusted CITS
Intercept (β_{00})	6.71 (0.26)*	8.65 (0.27)*	4.44 (0.47)*
Level 1			
[a]Pre TxSlope (β_{10})	0.21 (0.02)*	0.27 (0.03)*	0.34 (0.03)*
[a]TxLevel (β_{20})	0.34 (0.10)*	0.31 (0.12)*	0.29 (0.12)*
[a]Post TxSlope (β_{30})	0.02 (0.06)	0.02 (0.08)	-0.06 (0.08)
[a]Enroll (β_{40})			0.44 (0.03)*
White (β_{50})			-0.87 (0.41)*
[a]Poverty (β_{60})			-0.98 (0.26)*
Ratio (β_{70})			-0.00 (0.00)
Level 2			
TxGroup (β_{01})		-6.67 (0.50)*	-3.64 (0.34)*
TxGroup*Pre TxSlope (β_{11})		-0.20 (0.05)*	-0.22 (0.05)*
TxGroup*TxLevel (β_{21})		0.09 (0.22)	0.14 (0.22)
TxGroup*Post TxSlope (β_{31})		0.01 (0.14)	0.02 (0.14)

Note. *$p < .05$; [a] random effects; standard error in parentheses; For Enroll, as enrollment increases by 100 students AP subject offerings increases by .43 subjects

Covariate-Adjusted CITS

We added poverty, racial/ethnic composition, enrollment, and student-teacher ratio as time-varying Level-1 covariates in the covariate-adjusted CITS model. Inclusion of the time-varying covariates allowed us to determine if shifts in school demographics were responsible for the initial increase in AP course offerings observed in the CITS model. The last column of Table 3 presents results for the covariate-adjusted CITS. Although smaller schools, higher poverty schools and those with a higher percentage of White students had relatively lower AP course offerings, the statistical significance and direction of AP incentive effects did not change with the addition of the time-varying demographic covariates. That is, the covariate-adjusted CITS still revealed that the introduction of the AAC indicator resulted in an initial but not sustained increase in AP course offerings across all schools, with no significant difference between policy sensitive and nonsensitive schools. Figure 1 below presents the central findings from this study using the CITS model since the actual data and model results correspond so closely.

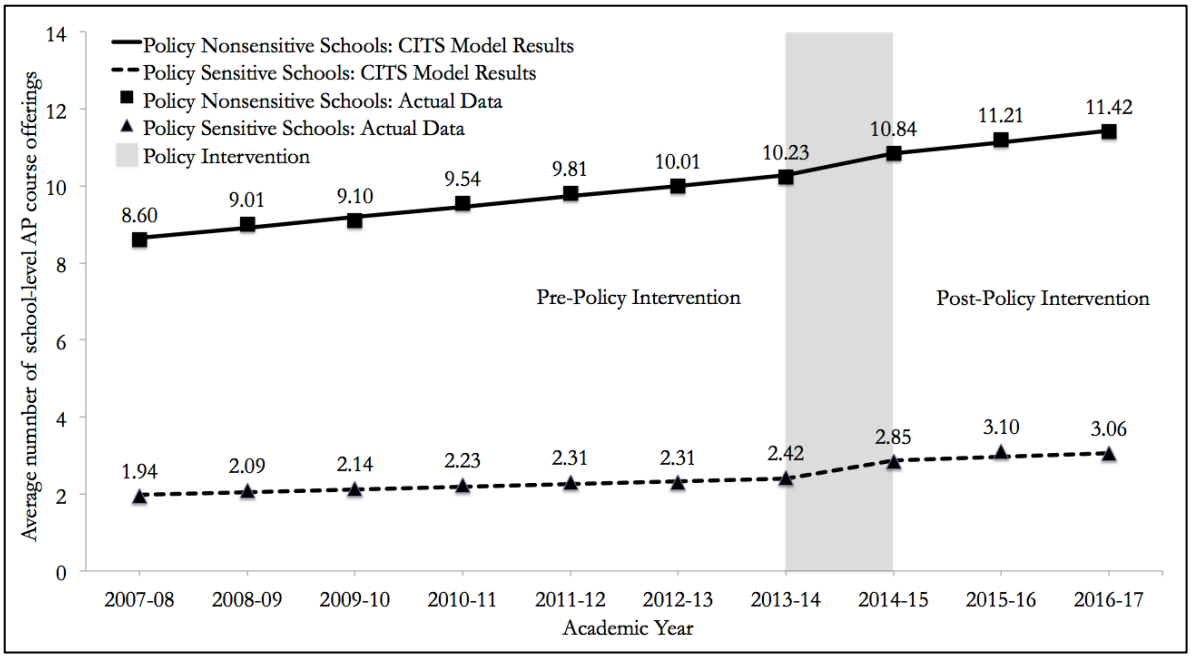

Figure 1. AP course offerings as functions of time and school group in the CITS model

In Figure 1, the initial disparity in AP course offerings, the differential pre-policy growth trajectories between policy sensitive and nonsensitive schools, and the initial increase in AP courses for both school groups coincident with the AP policy intervention can clearly be seen. The disparity in initial AP course offerings is represented by the large gap between policy sensitive and nonsensitive schools in academic year 2007-08. The differential AP growth trajectories are marked by the widening gap between policy sensitive and nonsensitive schools across time. Finally, the statistically significant TxLevel finding is evident in the noticeable shift in average AP course offerings for both school groups from 2013-14 (the academic year before Pennsylvania introduced the AAC indicator) to 2014-15 (the first policy intervention year).

Discussion

Despite increased public interest in AP and state-level policies designed to improve access (Education Commission of the States, 2016), research on the effects of those policies is relatively rare. Our study explored the effect of using accountability, not resource, incentives to improve school-level AP access. We focused on Pennsylvania's accountability reforms to examine if including an AAC indicator improved school-level AP access. Because Pennsylvania's AAC indicator resulted in two distinct school groups (policy sensitive and nonsensitive), our analyses enabled tests of whether the AAC indicator affected school groups differentially and if the two school groups differed demographically. Our aim was to analyze whether Pennsylvania's school accountability incentive had the intended effect of improving AP access.

Key Findings

Pennsylvania's inclusion of an AAC indicator in its accountability system led to an initial increase in AP course offerings, on average. The increase was short-lived, however, as the AAC indicator failed to alter long-term growth in school-level AP course offerings across schools. Stated differently, we found no evidence that the AAC indicator changed the rate at which schools increased or decreased AP course offerings relative to the baseline period. Our results also suggest

the AAC indicator influenced policy sensitive and nonsensitive schools similarly: both school groups increased AP course offerings immediately after the policy intervention, but then returned to pre-policy intervention trajectories. Moreover, the covariate-adjusted CITS showed the statistical significance and direction of these policy intervention effects did not change when we controlled for changes in school demographics across time.

Although we found no differences in how the two school groups reacted to the AAC indicator, our results revealed that an already sizeable gap in AP course offerings between policy sensitive and nonsensitive schools widened from 6.65 to 8.36 between academic years 2007-08 and 2016-17 (see Figure 1). This finding suggests that Pennsylvania's AAC indicator did not disrupt access inequalities between schools with the fewest and the most AP course offerings, supporting Klugman's (2013) analysis of three California policies meant to improve AP access. That is, the present study provides additional evidence for the theory of Effectively Maintained Inequalities, which asserts that privileged schools will pursue status enhancements associated with expansive AP course offerings to maintain comparative advantages over less privileged schools, often widening opportunity gaps. Klugman's study focused on a collection of resource-based policies; the present study showed similar findings when a state's accountability system relies on public pressure rather than resource investment.

Though not the primary objective of this study, our analyses revealed demographic differences between policy sensitive and nonsensitive schools. Those differences point to factors that associate with AP access inequalities and, therefore, might help formulate state and local policies that can leverage effective and equitable change. To that end, our comparisons of school groups in Pennsylvania echoed findings about unequal opportunities in California, Florida, Ohio, and various other regions of the U.S. (Cucchiara & Horvat, 2014; Demareth, 2009; Klugman, 2013; Oakes, Wells, Jones, & Datnow, 1997; Perna et al., 2015; Zarate & Pachon, 2006). Policy sensitive schools in Pennsylvania featured higher concentrations of students from socioeconomically disadvantaged backgrounds and students of color, highlighting troubling opportunity gaps that seem to be conditioned on socioeconomic status and race/ethnicity.

In addition to the aforementioned demographic differences, policy sensitive schools had smaller enrollments, supporting the notion that schools' abilities to exploit economies of scale also seems to matter for AP access. Perhaps reflecting a link between geographic locale and enrollment, policy sensitive schools were also more urban and more remote than policy nonsensitive schools. Accordingly, suburban schools seem better positioned to provide AP access, thus perpetuating inequalities elsewhere (see Rury & Saatcioglu, 2011). The geographic locale findings are also pertinent to Klugman's (2013) warning about the dangers of "potentially neglecting important dimensions of inequality" (p. 8). He referenced optimal ways to account for student- and school-level poverty indicators. Extending his argument to focus on nuances in geographic locale, another potential source of inequality, we remind other researchers about the importance of accounting for geography's complexity (Thier et al., 2017). Students living further away from cities (i.e., distant or remote) were most likely clustered into policy sensitive schools, where AP opportunities were most limited. Students in suburbs or fringes of those suburbs—which seem to be distinctions without any difference (Greenough & Nelson, 2015)—enjoyed greater AP access than peers in the innermost urban and outermost rural areas.

Policy Implications

Since the 1990s, school accountability has arguably been the most visible state policy approach for improving public schools (Shepard, 2008). The theory of change supporting school accountability incentives is that public reporting of school data will generate public pressure that causes teachers and administrators to change their behaviors leading to improved school

performance (Bae, 2018; Figlio & Loeb, 2011; Hess, 2002; Lee & Reeves, 2012). Some scholars describe this theory of change as "fatally simple" (Lee & Reeves, 2012, p. 210), while others argue high stakes school accountability systems can have "transformative" effects on students and educators through "the coercive force of self-interest" (Hess, 2002, p. 70). Neither argument fits neatly with our results, which show Pennsylvania's school accountability incentive resulted in an initial, but not sustained increase in AP course offerings. That is, Pennsylvania's decision to publicly report AP access data within an accountability framework had neither simple nor transformative effects, and was not a long-term solution for improving AP access across all schools or narrowing disparities in access between schools.

Two strands of principal-agent theory produce a possible explanation for the short-term gains achieved by Pennsylvania schools, but the lack of long-term, sustained increases in AP access. First, principal-agent theory maintains the strength of incentives can dictate the type of improvement efforts individuals or organizations pursue (Milgrom & Roberts, 1988). For example, research shows accountability systems that institute severe sanctions for poor performance incentivize schools to pursue short-term gains instead of investing in long-term solutions (Figlio & Loeb, 2011). Importantly, the severity of consequences attached to accountability indicators varies widely both within and between states (Polikoff et al., 2014). At the most severe end of the spectrum, some states use the information generated from individual accountability indicators both in the calculation of school quality ratings and to identify low-performing schools that will receive state mandated interventions. States with the weakest incentives simply report school performance data, but do not use these data in school quality rating calculations or to identify schools in need of improvement. In general, Pennsylvania's AP accountability incentive falls in the middle of this spectrum since the AAC indicator is used to calculate school quality ratings but not to identify schools in need of improvement (Polikoff et al., 2014). However, the strength of the incentive was weaker for policy sensitive schools since these schools, which maxed out on AAC points, could not earn more accountability points from increasing AP course offerings. Regardless, one possible explanation for the short-term increase in AP course access is that the strength of the incentive created by Pennsylvania's AAC indicator resulted in all schools improving AP access in the short-term above and beyond what the pre-policy intervention growth trend projected.

Second, lacking the necessary capacity to change or having inadequate resources to do so are possible explanations for a lack of long-term positive outcomes stemming from school accountability incentives (Figlio & Ladd, 2015). Stated differently, schools simply may not have had the capacity or resources to sustain significant increases in AP course offerings. From a capacity perspective, several researchers suggest the challenge of identifying a sufficient number of students with high prior achievement is a barrier that may prevent educators from creating or demanding new AP courses for interested students (e.g., Iatarola et al., 2011; Kolluri, 2018; Rowland & Shircliffe, 2016). Referring to this situation as "gatekeeping," Rowland & Shircliffe (2016) highlight the fears of some educators that students with low levels of prior achievement will not benefit from AP participation, while also potentially interfering with experiences of more able students who sought early opportunities to learn college-level content and potentially earn college credit.

Having a critical mass of students with high prior achievement to demand AP is just one resource-based explanation for why large schools can offer AP courses more readily than small schools. Another large school advantage includes the likelihood of having more teachers with advanced degrees and simply having more teachers in general (Iatarola et al., 2011). Challenges of recruiting and retaining highly qualified teachers also differentially disadvantage high-poverty schools (Clotfelter, Ladd, Vigdor, & Wheeler, 2007; Simon & Johnson, 2015) and schools in rural areas (Azano & Stewart, 2015). Policy sensitive schools in our sample were partly defined by being both

high-poverty and in remote areas. This may be one reason the gap between these schools and policy nonsensitive schools widened across time.

Relying on only capacity limitations or resource inadequacy to explain disparities in school-level AP access, however, does not pay sufficient attention to the structural, social, and political barriers that prevent equal access to advanced coursework to students in schools serving high proportions of students from socioeconomically disadvantaged backgrounds or students of color (e.g., academic tracking based on race/ethnicity or class; see Oakes, 1985; 2005). Reflecting on 30 years of research on academic tracking, Oakes (2018) maintains that community fears attached to racial integration, a belief in a fixed intelligence for individuals, and the "politics of comparative advantage" has thwarted the effectiveness and sustainability of most detracking policies and programs (p. 96). In other words, structural, social, and political barriers may be just as important to explaining between-school disparities in AP access as the constraints that resource deprivation and geography present. Notably, the theory of change implicit in Pennsylvania's accountability system design does not directly address the potential root causes to AP access disparities associated with capacity limitations, resource inadequacy, or the barriers posed by academic tracking.

Taken together, these two strands of principal-agent theory suggest schools reacted to the moderate severity of the consequences attached to the AAC indicator by pursuing short-term gains. However, schools either lacked the capacity, had inadequate resources, or confronted other barriers that prevented the sustainability of those gains. Therefore, it appears comprehensive AP accountability polices should pay sufficient attention to the strength of the incentives produced by the design of individual indicators as well as invest in building the capacity of schools to increase AP access, address resource inadequacy, and challenge the harmful norms that perpetuate academic tracking based on student demographics.

Limitations

This study asked whether the introduction of Pennsylvania's AP accountability incentive significantly changed school-level AP access immediately and across time. Our results suggest the initial answer to that question is yes in the short-term but no in the long-term. We acknowledge, however, that our conclusion rests on evidence from one state. Variation among state approaches to AP accountability incentives involve several types of indicators, weights placed on those indicators, and potential overlaps with other incentive policies and programs. Further studies that employ quasi-experimental designs or rigorous alternatives are needed to examine the effects of other state AP accountability incentives.

Moreover, conducting ITS designs require researchers' attentiveness to internal and external validity threats (see Hallberg, Williams, Swanlund, & Eno, 2018; Shadish et al., 2002; Wong, Cook, & Steiner, 2015). The internal threat of *history* centers on whether separate, alternative occurrences—explicit policies or external events—could explain the presence or absence of intervention effects. To represent a true threat to validity, separate, alternative occurrences must be coincident with the onset of policy intervention. For example, although Pennsylvania features few state AP policies in comparison to other states, expanding AP access is one of 12 approved purposes for which school districts can use funds from the state's Ready to Learn block grant program. In academic year 2014-15, the Ready to Learn program replaced the Accountability Block Grant program, which began in 2004-05. In addition to renaming the program, the state increased Ready to Learn funding by $100 million in 2014-15, the year after Pennsylvania launched the *School Performance Profile*. Ready to Learn funding also increased from $200 million in 2014-15 to $250 million in 2016-17. Since changes to the Ready to Learn program occurred after Pennsylvania introduced the *School Performance Profile,* this potential threat to our findings was in relation to the post-policy intervention slope, for which we found no statistically significant association.

Readers of the present study should, however, cautiously interpret the findings from this study because of two external validity threats: interactions of the casual relations with treatment variations and settings (see Shadish et al., 2002). First, because AP accountability incentives can manifest in many ways, these findings are not generalizable to *all* AP accountability incentives. For instance, Pennsylvania's AAC indicator also included International Baccalaureate and other college credit-bearing coursework, potentially weakening the power of the AP accountability incentive, particularly for schools with robust International Baccalaureate or similar programs that provide access to college-level coursework. Though in Pennsylvania, AP far outweighs International Baccalaureate programming as only 16 public schools offer the International Baccalaureate Diploma Programme, designed for high school aged students (International Baccalaureate Organization, 2018). Second, findings from this study are limited to only traditional high schools in Pennsylvania that were rated on the *School Performance Profile*. Findings may vary when examining different types of schools in different states that operate under different types of accountability systems. As the Education Commission of the States 2016 report shows, the design of AP accountability incentives varies considerably across states.

Conclusion

Writing 15 years ago, Klopfenstein (2004) noted that "[i]ncreasing amounts of money are being invested, privately and at all levels of government, to expand the AP Program nationwide, but there has been little detailed research regarding the effect of this money on the access and participation of traditionally underserved students" (p. 10). Unfortunately, not much has changed. We still know very little about what policies will effectively expand AP access in schools, especially if those schools predominantly serve students from socioeconomically disadvantaged backgrounds and/or students of color, if those schools are small, or if those schools reside outside suburbs. Furthermore, as Kolluri's (2018) literature review demonstrates, disparities in AP access between schools remain despite several federal and state policy initiatives.

As long as some students have their pathways impeded from the trajectory-altering opportunity of college-level coursework during high school (Adelman, 1999; 2006; Engberg & Wolniak, 2010), state educational systems will have to confront associated problems such as inequitable college admissions, on-time graduation, and earned income. For example, one benefit of AP courses is the possibility that a student could score high enough on end-of-course exams to bypass some college prerequisite courses. The resulting flexibility in collegiate scheduling could facilitate early or on-time graduation in an era when only about 60% of full-time degree-seeking students earn degrees within six years (National Center for Education Statistics, 2017). Consequently, one should not just consider additional tuition fees associated with expenditures beyond the four-year proposition that most incoming college students expect. One should also factor opportunity costs of lost income when completing four-year degrees takes, on average, 50% longer than most might have planned for. To alleviate those concerns, we call for examinations in other states to determine if, and under what conditions, AP accountability incentives narrow access disparities. In the absence of such research, state policymakers will continue creating school accountability incentives with little idea of what actually works.

References

Achieve, & Jobs for the Future. (2015). *Integrating earning college credit in high school into accountability systems*. Washington, DC & Boston, MA: Author.

Ackerman, P. L., Kanfer, R., & Calderwood, C. (2013). High school Advanced Placement and student performance in college: STEM majors, non-stem majors, and gender differences. *Teachers College Record*, *115*(10), 1–43.

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment*. Washington, DC: U.S. Department of Education.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.

Azano, A. P., & Stewart, T. T. (2015). Exploring place and practicing justice: Preparing pre-service teachers for success in rural schools. *Journal of Research in Rural Education*, *30*(9), 1–12.

Bae, S. (2018). Redesigning systems of school accountability: A multiple measures approach to accountability and support. *Education Policy Analysis Archives*, *26*(8), 1–32. https://doi.org/10.14507/epaa.26.2920

Barnard-Brak, L., McGaha-Garnett, V., & Burley, H. (2011). Advanced Placement course enrollment and school-level characteristics. *NASSP Bulletin*, *95*, 165–174. https://doi.org/10.1177/0192636511418640

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4* (R package version 1.1-7) [Computer software]. https://doi.org/10.18637/jss.v067.i01

Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, *41*, 287–302. https://doi.org/10.1080/0305764X.2011.607151

Cisneros, J., Gomez, L. M., Powers, J. M., Holloway-Libell, J., & Corley, K. M. (2014). The Advanced Placement opportunity gap in Arizona: Access, participation, and success. *AASA Journal of Scholarship & Practice*, *11*, 20–33.

Clotfelter, C., Ladd, H., Vigdor, J., & Wheeler, J. (2007). *High-poverty schools and the distribution of teachers and principals*. Chapel Hill, NC: North Carolina Law Review, University of North Carolina. https://doi.org/10.1037/e722752011-001

College Board. (1997). *1997 Advanced Placement program national summary reports*. Retrieved from http://media.collegeboard.com/digitalServices/pdf/research/national_1997.pdf

College Board. (2014). *The 10th annual AP report to the nation*. Author: New York, NY. Retrieved from http://media.collegeboard.com/digitalServices/pdf/ap/rtn/10th-annual/10th-annual-ap-report-to-the-nation-single-page.pdf

College Board. (2016). *National Report*. Retrieved from https://research.collegeboard.org/programs/ap/data/archived/ap-2016

College Board. (2017). *National Report*. Retrieved from https://research.collegeboard.org/programs/ap/data/archived/ap-2017

College Board. (2018). *AP® Program Guide 2018-19*. Retrieved from https://apcentral.collegeboard.org/pdf/ap-program-guide-2018-19.pdf

Conger, D., Long, M. C., & Iatarola, P. (2009). Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management*, *28*, 555–576. https://doi.org/10.1002/pam.20455

Conley, D. T., Thier, M., Beach, P., Lench, S. C., & Chadwick, K. L. (2014). *Measures for a college and career indicator: Advanced coursework*. Eugene, OR: Educational Policy Improvement Center.

Cucchiara, M. B., & Horvat, E. M. (2014). Choosing selves: The salience of parental identity in the school choice process. *Journal of Education Policy, 29*, 486–509. https://doi.org/10.1080/02680939.2013.849760

Darity, Jr., W., Castellino, D., Tyson, K., Cobb, C., & McMillen, B. (2001). *Increasing opportunity to learn via access to rigorous courses and programs: One strategy for closing the achievement gap for at-risk and*

ethnic minority students. Raleigh, NC: Division of Accountability, North Carolina State Department of Public Instruction.

Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives*, *22*(86), 1–37. https://doi.org/10.14507/epaa.v22n86.2014

Demareth, P. (2009). *Producing success: The culture of personal advancement in an American high school*. Chicago, IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226142425.001.0001

Education Commission of the States. (2006). *Advanced Placement policies: All state profiles*. Retrieved from http://ecs.force.com/mbdata/MBQuestRT?Rep=AP03

Education Commission of the States. (2016). *Advanced Placement policies: All state profiles*. Retrieved from http://ecs.force.com/mbdata/mbprofallrt?Rep=APA16

Education Trust. (2019). *Advanced coursework: An overview*. Washington, DC: Author.

Engberg, M. E., & Wolniak, G. C. (2010). Examining the effects of high school contexts on postsecondary enrollment. *Research in Higher Education*, *51*, 132–153. https://doi.org/10.1007/s11162-009-9150-y

Figlio, D., & Ladd, H. F. (2015). School accountability and student achievement. In H. F. Ladd & M. Goertz (Eds.), *Handbook of research in education finance and policy* (2nd ed., pp. 194–210). New York: Routledge.

Figlio, D., & Loeb, S. (2011). School accountability. In E. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 3, pp. 383–421). North Holland, The Netherlands: Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00008-9

Gagnon, D. J., & Mattingly, M. J. (2016). Advanced Placement and rural schools: Access, success, and exploring alternatives. *Journal of Advanced Academics*, *27*, 266–284. https://doi.org/10.1177/1932202X16656390

Geiser, S., & Santelices, V. (2004). *The role of Advanced Placement and honors courses in college admissions*. Berkeley, CA: Center for Studies in Higher Education, University of California Berkeley.

Greenough, R., & Nelson, S. R. (2015). Recognizing the variety of rural schools. *Peabody Journal of Education*, *90*, 322–332. https://doi.org/10.1080/0161956X.2015.1022393

Hallberg, K., Williams, R., Swanlund, A., & Eno, J. (2018). Short comparative interrupted time series using aggregate school-level data in education research. *Educational Researcher*, *47*, 295–306. https://doi.org/10.3102/0013189X18769302

Hess, F. M. (2002). Reform, resistance... retreat? The predictable politics of accountability in Virginia. In D. Ravitch (Ed.)*, Brookings Papers on Educational Policy: 2002* (pp. 69-122). Washington, DC: Brookings Institution Press. https://doi.org/10.1353/pep.2002.0007

Holmstrom, B., & Costa, J. R. (1986). Managerial incentives and capital management. *The Quarterly Journal of Economics*, *101*, 835–860. https://doi.org/10.2307/1884180

Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, *7*, 24–52. https://doi.org/10.1093/jleo/7.special_issue.24

Iatarola, P., Conger, D., & Long, M. C. (2011). Determinants of high schools' advanced course offerings. *Educational Evaluation and Policy Analysis*, *33*, 340–359. https://doi.org/10.3102/0162373711398124

International Baccalaureate Organization. (2018). Find an IB World School. Retrieved from https://www.ibo.org/programmes/find-an-ib-school/

Jackson, C. K. (2010). A little now for a lot later: A look at a Texas Advanced Placement Incentive Program. *Journal of Human Resources*, *45*, 591–639. https://doi.org/10.1353/jhr.2010.0019

Jensen, M. C., & Meckling, W. C. (1976). Theory of the firm: Managerial behavior, agency costs and

ownership structure. *Journal of Financial Economics*, *3*, 305–360. https://doi.org/10.1016/0304-405X(76)90026-X

Jeong, D. W. (2009). Student participation and performance on Advanced Placement exams: Do state-sponsored incentives make a difference? *Educational Evaluation and Policy Analysis*, *31*, 346–366. https://doi.org/10.3102/0162373709342466

Kramer, D. (2016). Examining the impact of state level merit-aid policies on Advanced Placement participation. *Journal of Education Finance*, *41*, 322–343. https://doi.org/10.1353/jef.2016.0000

Klopfenstein, K. (2004). The Advanced Placement expansion of the 1990s: How did traditionally underserved students fare? *Education Policy Analysis Archives*, *12*(68), 1–15. https://doi.org/10.14507/epaa.v12n68.2004

Klugman, J. (2013). The Advanced Placement arms race and the reproduction of educational inequality. *Teachers College Record*, *115*(5), 1–34.

Kolluri, S. (2018). Advanced Placement: The dual challenge of equal access and effectiveness. *Review of Educational Research*, *88*, 671–711. https://doi.org/10.3102/0034654318787268

Lacy, T. (2010). Access, rigor, and revenue in the history of the Advanced Placement program. In P. Sadler, G. Sonnert, R. Tai, & K. Klopfenstein (Eds.), *A critical examination of the Advanced Placement program* (pp. 17–50). Cambridge, MA: Harvard Education Press.

Lee, J., & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990–2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis*, *34*, 209–231. https://doi.org/10.3102/0162373711431604

Lucas, S. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. *American Journal of Sociology*, *106*, 1642–1690. https://doi.org/10.1086/321300

Malkus, N. (2015). *AP at scale: Public school students in Advanced Placement, 1990-2013*. Washington, DC: American Enterprise Institute.

Malkus, N. (2016). *The AP peak: Public school offerings Advanced Placement, 2000-2012*. Washington, DC: American Enterprise Institute.

Milgrom, P., & Roberts, J. (1988). An economic approach to influence activities in organizations. *American Journal of Sociology*, *94*, S154-S179. https://doi.org/10.1086/228945

Moe, T. M. (2003). The politics and practice of accountability. In P. E. Peterson & M. R. West (Eds.), *No Child Left Behind? The politics and practice of school accountability*. Washington D.C.: Brookings Institution Press.

Moe, T. M. (1984). The new economics of organization. *American Journal of Political Science*, *28*, 739–777. https://doi.org/10.2307/2110997

Morse, R., & Brooks, E. (2019). How U.S. News calculated the 2019 Best High School ranking. *U.S. News & World Report*. Retrieved from https://www.usnews.com/education/best-high-schools/articles/how-us-news-calculated-the-rankings

National Center for Education Statistics. (2017). Graduation rate from first institution attended for first-time, full-time bachelor's degree-seeking students at 4-year postsecondary institutions, by race/ethnicity, time to completion, sex, control of institution, and acceptance rate: Selected cohort entry years, 1996 through 2010. Retrieved from https://nces.ed.gov/programs/digest/d17/tables/dt17_326.10.asp?current=yes

Oakes, J. (1985). *Keeping track*. New Haven, CT: Yale University Press.

Oakes, J. (2005). *Keeping track* (2nd ed.). New Haven, CT: Yale University.

Oakes, J. (2018). 2016 AERA presidential address: Public scholarship: Education research for a diverse democracy. *Educational Researcher*, *47*, 91–104. https://doi.org/10.3102/0013189X17746402

Oakes, J., Wells, A. S., Jones, M., & Datnow, A. (1997). Detracking: The social construction of ability, cultural politics, and resistance to reform, *Teachers College Record*, *98*, 482–510.

Owens, A., Reardon, S., & Jencks, C. (2016). Income segregation between schools and school districts. *American Educational Research Journal*, *53*, 1159–1197. https://doi.org/10.3102/0002831216652722

Perna, L. W., May, H., Yee, A., Ransom, T., Rodriguez, A., & Fester, R. (2015). Unequal access to rigorous high school curricula: An exploration of the opportunity to benefit from the International Baccalaureate Diploma Programme. *Educational Policy*, *29*, 402–425. https://doi.org/10.1177/0895904813492383

Polikoff, M. S., McEachin, A. J., Wrabel, S. L., & Duque, M. (2014). The waive of the future? School accountability in the waiver era. *Educational Researcher*, *43*, 45–54. https://doi.org/10.3102/0013189X13517137

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage

Rothstein, R., Jacobsen, R., & Wilder, T. (2008). *Grading education: Getting accountability right*. Washington, DC and New York: Economic Policy Institute and Teachers College Press.

Rowland, M. L., & Shircliffe, B. J. (2016). Confronting the "acid test:" Educators' perspectives on expanding access to Advanced Placement at a diverse Florida high school. *Peabody Journal of Education*, *91*, 404–420. https://doi.org/10.1080/0161956X.2016.1184947

Rury, J. L., & Saatcioglu, A. (2011). Suburban advantage: Opportunity hoarding and secondary attainment in the postwar metropolitan North. *American Journal of Education*, *117*, 307–342. https://doi.org/10.1086/659210

Sadler, P. M. (2010). Advanced high school coursework and college admission decisions. In P. Sadler, G. Sonnert, R. Tai, & K. Klopfenstein (Eds.), *A critical examination of the Advanced Placement program* (pp. 245–261). Cambridge, MA: Harvard Education Press.

Schneider, J. (2009). Privilege, equity, and the Advanced Placement program: Tug of war. *Journal of Curriculum Studies*, *41*, 813–831. https://doi.org/10.1080/00220270802713613

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage learning.

Shepard, L. A. (2008). A brief history of accountability testing, 1965-2007. In K. E. Ryan, & L. A. Shepard (Eds.), *The future of test-based educational accountability* (pp. 25–46). New York: Routledge.

Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195152968.001.0001

Simon, N. S., & Johnson, S. M. (2015). Teacher turnover in high-poverty schools: What we know and can do. *Teachers College Record*, *117*(3), 1–36.

Thier,, M., Beach, P., Martinez, C. R., & Hollenbeck, K. L. (2017, April). *Take care when cutting: Unintended consequences of disaggregating school-level data as rural*. Paper presented at the annual meeting of the American Education Research Association, San Antonio, TX.

Thier, M., Beach, P., Todd, J., & Coleman, M. (2016). Places, poverty, and pupils: Exploring key predictors of Advanced Placement access. Paper presented at the annual meeting of the American Education Research Association, Washington, DC.

U.S. Department of Education. (2013). *Advanced Placement incentive program grants: Funding status*. Retrieved from https://www2.ed.gov/programs/apincent/funding.html

Warne, R. (2017). Research on the academic benefits of the Advanced Placement program: Taking stock and looking forward. *SAGE Open*, *7*(1), 1–16. https://doi.org/10.1177/2158244016682996

Wong, M., Cook, T. D., & Steiner, P. M. (2015). Adding design elements to improve time series designs: No Child Left Behind as an example of causal pattern-matching. *Journal of Research on Educational Effectiveness*, *8*, 245-279.

Zarate, M., & Pachon, H. (2006). *Equity in offering advanced placement courses in California high schools.* Los Angeles, CA: The Tomás Rivera Policy Institute.

About the Authors

Paul Beach
University of Oregon
ptb@uoregon.edu
https://orcid.org/0000-0001-8121-2459
Paul Beach is a PhD candidate in Quantitative Research Methods in Education at the University of Oregon and a Research Associate at Inflexion. His research focuses on analyzing the effects of state accountability policy and the factors that influence equitable school improvement.

Keith Zvoch
University of Oregon
kzvoch@uoregon.edu
http://orcid.org/0000-0001-9348-5155
Keith Zvoch is an Associate Professor in the Department of Educational Methodology, Policy and Leadership at the University of Oregon. His research focuses on causal inference in field-based settings, the application and study of methodologies that account for breakdowns in treatment protocol, and the modeling of time series data.

Michael Thier
University of Oregon
mthier@uoregon.edu
http://orcid.org/0000-0003-2851-6255
Michael Thier is a PhD candidate in Educational Leadership at the University of Oregon. His research focuses on global citizenship education, including international/cross-cultural comparisons and the opportunities and conditions that enable global citizenship education in rural and remote settings.

education policy analysis archives

Volume 27 Number 138 October 28, 2019 ISSN 1068-2341

Please send errata notes to Audrey Amrein-Beardsley at audrey.beardsley@asu.edu

Join EPAA's Facebook community at https://www.facebook.com/EPAAAAPE and **Twitter feed** @epaa_aape.