Title

Stratification through Organizational Identity: Charter School Ideologies and Segregation in the Era of Accountability

Author

Haber, Jaren

Publication Date

2020

Peer reviewed|Thesis/dissertation

Stratification through Organizational Identity:
Charter School Ideologies and Segregation in the Era of Accountability

by

Jaren Haber

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy
in
Sociology

in the
Graduate Division
of the

University of California, Berkeley

Committee in charge:

Professor Heather Haveman, Chair
Professor Samuel Lucas
Professor Calvin Morrill

Spring 2020

ABSTRACT

Stratification through Organizational Identity:
Charter School Ideologies and Segregation in the Era of Accountability

by

Jaren Haber

Doctor of Philosophy in Sociology

University of California, Berkeley

Professor Heather Haveman, Chair

Research shows charter schools are more segregated by race and class than traditional public schools. I investigate an under-examined mechanism for this segregation: Charter schools project identities corresponding to parents' race- and class-specific parenting styles and educational values. Such identities are constituted by conformity and uniqueness claims reflecting cultural logics—especially the ascendant logic of school accountability, which compels standardization despite pressures on charter schools for locally situated innovation.

To analyze these tensions, I use computational text analysis to detect emphasis on educational ideologies in the websites of all charter schools operating in 2015-16. In particular, I use Structural Topic Models to discover a range of latent themes; I then use word embeddings and dictionary methods to measure inquiry-based learning (IBL), a particularly niche-specific ideology. Finally, I implement mixed linear regression models to test the relationships between ideological emphasis and school- and district-level poverty and ethnicity. I thereby transcend methodological problems in scholarship on charter school identities by collecting contemporary, population-wide data, and by blending text analysis with hypothesis testing.

I find that the websites of schools serving white and affluent students and districts are more likely to emphasize academics (e.g., course information, IBL), while schools serving those in poverty or people of color place greater emphasis on standards, college preparation, and programs and services for the disadvantaged (e.g., nutrition, virtual learning). For IBL, precise statistical models show that these relationships are robust, even when objective measures of school quality are considered. These findings suggest charter school identities are both race- and class-specific, outlining a new mechanism by which school choice may consolidate parents by race and class—and paving the way for behavioral and longitudinal studies. This dissertation contributes to literatures on school choice, quantitative methods, and educational stratification.

ACKNOWLEDGMENTS

I am especially grateful to Heather Haveman for her constructive criticism, which after many drafts have greatly improved this dissertation. I also give thanks to the UC Berkeley D-Lab community for teaching me to code and to embrace not knowing. I acknowledge also Sam Lucas, Calvin Morrill, Bruce Fuller, David Bamman, Ben Gebre-Medhin, and Caroline Le Pennec-Caldichoury for their feedback and insightful comments; Aaron Culich, Carl Mason, and the Cloud Working Group for help with web data collection and computing infrastructure; Jae Yeon Kim for help with visualization; and my family and wife-to-be for their encouragement and support.

This complex project wouldn't have been possible without the contributions of 38 research assistants from UC Berkeley's Undergraduate Research Apprentice Program and Data Science Discovery Program. Listed alphabetically, they are: Kanika Ahluwalia, Brad Afzali, Akcan Balkir, Muying Chen, Siyuan Chen, Yitong Chen, Kaan Dogusoy, Saabhir Gill, Harshayu Girase, Akshat Gokhale, Yoon Sung Hong, Jennifer Huang, Elaine Huynh, Krutika Ingale, Jiyoon Jeong, James Jung, Inderpal Kaur, Francis Kumar, Yong Jin Kweon, Ariel Langer, Brian Yimin Lei, Xueyong Liu, Haley Miller, Anna Nguyen, Thao Nguyen, Tina Nguyen, Madeleine Peng, Emily Qian, Samyukta Raman, Ji Shi, Sarah Solieman, Arjun Srinivasan, Prianka Subrahmanyam, Frank Wang, Violet Yao, George Wu, Max Yuan, and Jiahua Zou.

Previous versions of this project and its methods were presented at the Berkeley Institute for Data Science's 2018 Text Across Domains (TextXD) symposium; the D-Lab's Computational Text Analysis Working Group in 2017-18; the 2018 Making Text Research-Ready symposium; the 2018 Graduate School of Education Research Day; and the American Sociology Association's Sociology of Education Section in 2017 and 2018. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562, as well as the Berkeley Demography Lab cloud computing facility. Financial support was provided by the UC Berkeley Dissertation Completion Fellowship and the Bridge Lowenthal Fellowship.

REPRODUCIBILITY

Condensed portions of this dissertation are forthcoming in the journal *Sociology of Education* under the title "Sorting Schools: A Computational Analysis of Charter School Identities and Stratification". A post-print of that article is available at http://bit.ly/OSF-postprint-SocEd; replication data and code are available at http://bit.ly/sorting-schools-2019. A pre-registration of my study design is available at http://bit.ly/OSF-prereg-SocEd.

DATA ACCESSIBILITY

The original web-crawled data have not been publicly released pursuant to the copyrights some charter schools have on their web content. Data access has been granted only to research assistants for the duration of their position. Researchers interested in accessing the original web-crawling data may contact the author.

RESEARCH ETHICS

The research reported here used only publicly available, organization-level data and did not involve human subjects. As such, it was not subject to review by the institutional review board for the protection of human subjects. Web-crawling speeds were throttled to prevent server overload.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

As the bastion of school choice, the educational reform of charter schools has won bipartisan support and grown tremendously in recent decades (Berends 2015). Charter schools are publicly funded yet free from the labor regulations (e.g., teachers' unions, certification requirements), content restrictions (e.g., district-mandated textbooks and programs), and financial oversight (e.g., regular audits, district record-keeping systems) facing traditional public schools (e.g., Orfield, Gumus-Dawes, and Luce 2013)—granting them autonomy to innovate in pursuit of greater performance (e.g., Lubienski 2003). Recent years have seen charter schools more than triple in number and a nearly nine-fold increase in the number of students served, climbing from 1,542 schools with 349,714 students in 1999-2000 to reach 6,992 schools serving 3,008,106 students in 2016-17 (National Alliance for Public Charter Schools 2018). Thus, in under two decades, charter schools have grown from 0.74% of the U.S. K-12 student population in the fall of 1999 to 5.95% in fall 2016 (National Center for Education Statistics 2018). Given the mounting prevalence of charter schools in U.S. society, this project examines the relationships between charter schools and their local communities.

All schools depend for survival on local funding and pupils, but as schools of choice, this is especially true for charter schools. Charter schools must be sufficiently attractive that local parents opt not for their neighborhood traditional public school but instead for a charter school. Yet presenting an appealing picture to parents is no simple matter—it depends on local parents' positions in broader social hierarchies of socioeconomic status and race, which structure their goals and interests. That is, what parents want for their children depends on cultural and social factors linked to their social positions and resources: parents and schools prioritize different habits and skills for different social strata (e.g., Kohn 1969; Lareau 2011), and parents also want their children to be around peers "of quality" (e.g., Abdulkadiroglu et al. 2019; Rothstein 2006) and socially similar to them in dimensions like race and social status (e.g., Holme 2002; Shrum, Cheek, and Hunter 1988). Thus, each charter school must present a strategic *identity*—a cultural conception projected to external audiences, namely parents, teachers, and oversight agencies—that resonates with local parents' socially embedded inclinations.

As organizations, the survival of charter schools depends on attaining both legitimacy—that is, inclusion in established categories with proven results (e.g., Zuckerman 1999)—on one hand, and uniqueness or innovation relative to peers, on the other (e.g., King, Clemens, and Fry 2011). Thus, charter schools seek to satisfy expectations for both uniformity—by clustering into recognizable identities (King et al. 2011) or generalizing the language of their mission statements (Renzulli, Barr, and Paino 2015)—and uniqueness—by recombining standard elements (King et al. 2011) or narrowing the focus of their mission statements (Renzulli et al. 2015). Rather than falling on either extreme of the assimilation/differentiation polarity, then, charter schools seek to establish identities that are both familiar and strategically different (Brewer 1991).

Indeed, to fit a sociodemographic niche, charter schools must both match expectations for that category and differentiate from peers and other categories (e.g., Porac et al. 1995). As such, charter schools deploy not only ritualized symbols of effectiveness like academic proficiency scores (Meyer and Rowan 1977), but also their identity claims adapt to the social backgrounds of those to whom they wish to appeal (e.g., Fuller 2009; Lauen, Fuller, and Dauter 2015). Specifically, charter schools project identities corresponding to the educational ideologies and parenting styles of race- and class-specific audiences—a mechanism for segregation by race and class and the focus of this study.

Capturing both isomorphism and differentiation (Huerta and Zuckerman 2009) in charter schools' identity claims is challenging—and forestalled by research designs that sort charter schools into single, uniform categories (e.g., Malkus and Hatfield 2017; Renzulli et al. 2015). I overcome this limitation by using text-analytic methods to continuously measure fine-grained, culturally embedded linguistic cues that schools use to signal their identities and appeal to specific sociodemographic niches.

Investigating charter schools' identities within local social contexts helps answer pressing questions in the sociology of education. Given their relative lack of direct district control (e.g., Paino, Boylan, and Renzulli 2017) and need to carefully cultivate a locally attractive image in order to secure resources, charter schools offer an opportunity to analyze how schools signal identities to navigate local instantiations of broader social hierarchies. Moreover, following research into charter school identities (e.g., King et al. 2011; Renzulli et al. 2015) and the unequal impacts of charter schools (e.g., Frankenberg and Siegel-Hawley 2013; Lacireno-Paquet et al. 2002), I investigate the role that distinct organizational forms play in parents' self-sorting into segregated schools and schools' self-sorting into segregated neighborhoods.

I argue that charter schools encourage self-sorting by (in part) projecting identities corresponding to parents' race- and class-specific parenting styles and educational values. Specifically, I propose that charter schools present different identities to communities that are advantaged—relatively affluent and white—compared to those that are disadvantaged—relatively poor or people of color. To test this proposition, I analyze an educational *ideology* (a set of beliefs about the social world that motivates moral action; Pettigrew 1979:575) common in schools: *inquiry-based learning* (IBL).

Ideology contrasts with *ritual*: broad normative appeals not intended for a specific cultural or sociodemographic audience. Moreover, as a central ideology of progressive-style educational programs (e.g., Dewey 1938), IBL contrasts with traditionalist models (e.g., No Excuses; Thernstrom and Thernstrom 2004) that emphasize basic, testable knowledge and skills and are favored by the politics of school accountability. In other words, I conceptualize two axes: ritual/ideology, on the one hand; and (within ideology) progressivist/traditionalist, on the other. These dynamics are discussed in more detail in the next chapter.

IBL entails student observation of phenomena, first-hand scientific inference, and student-centered construction of knowledge (e.g., Bruner 1961; Steffe and Gale 1995). IBL is widespread in charter as well as traditional public schools (Waite 2019) and has inspired broad policy initiatives including the Common Core (e.g., Watanabe 2007) and Next Generation Science Standards (Achieve 2010), efforts fueled by lively scholarship in the learning sciences.[1] In practice, IBL-based approaches favor the privileged: Arts; Classical; International; Science, Technology, Engineering, and Math (STEM); or Progressive models—including student-centered and experiential approaches (Waite 2019)—generally enroll more white students and

[1] The pedagogy of IBL (e.g., "What guidance is essential for student learning?", "To what extent should students 'act like scientists'?"), its efficacy (e.g., "How much learning do students retain in long-term memory?"), and even its terminology (e.g., "What is 'minimally guided instruction?'"; Hake 2008; Kirschner, Sweller, and Clark 2006) have been rigorously debated for decades (for a recent review, see Lazonder and Harmsen 2016).

fewer low-income students or students of color than do nearby traditional public schools (Malkus and Hatfield 2017). This tendency of IBL to segregate social groups makes it an especially sharp tool for studying how ideologies separate race and class factions among charter schools.

But little research has explored how IBL or other educational ideologies are invoked in race- and class-differentiated settings (but see Malkus and Hatfield 2017). I fill this gap through fine-grained, socially embedded analysis of educational discourse. Specifically, I address the research question: What is the relationship between a charter school's race and class composition and its emphasis on IBL? I predict that charter schools present themselves to affluent and white communities in ways emphasizing IBL.

To operationalize this ideology and study its connections with race and class distributions in charter schools, I capitalize on new computational tools for measuring culture inductively (e.g., Bail 2014; Nelson 2017) alongside the deductive method of mixed-effects linear regression. Moreover, I redress methodological and theoretical oversights in the school choice literature by collecting contemporary, valid, and population-wide charter school data and analyzing it with complementary text-analytic and statistical methods. To power my computational workflow, I used the flexible, open-source Python 3 (Van Rossum and Drake 2011) and R (R Core Team 2018) languages in the Extreme Science and Engineering Discovery Environment (XSEDE; Towns et al. 2014)[2] as well as Stata 15 (StataCorp 2017b) in the Berkeley Demography Lab cloud computing facility (see https://lab.demog.berkeley.edu/).

Specifically, I gathered rich, nationally comprehensive data on identities for all U.S. charter schools drawn from their websites. Websites are the collective hubs of many modern forms of organization, enabling collective action, information sharing, and new forms of interaction (e.g., Bennett and Segerberg 2013; Ferraro and O'Mahony 2012). Moreover, websites are rich, ecologically valid sources of cultural information targeting multiple audiences and revealing the organization's identity and goals (Powell, Horvath, and Brandtner 2016). As such, parents commonly rely on websites for signals on school quality and "fit" with their children—and in some cases, website information influences parents' perceptions of school quality more than do test scores (Yettick 2016).

The rest of this dissertation proceeds as follows. The next chapter sets the stage for my empirical analysis by briefly reviewing research on charter schools, inequality, and educational values and styles; describing the specific educational ideology I deductively test; and offering a hypothesis. In chapter 3, I ground my study in a theoretical framework built on institutional analysis, educational politics, and organizational identity. In chapter 4, I outline my data collection—with web-crawling and large-scale educational data sources—and computational methods: Structural Topic Modeling, dictionary approaches, word embeddings, and mixed linear regressions. In chapter 5, I describe the results of the Structural Topic Models, including by inductive clustering of topics based on their association with sociodemographics. In chapter 6, I narrow down to a single topic—inquiry-based learning—when describing the more robust results of dictionaries and mixed linear regressions. I discuss conclusions in the final chapter.

[2] In particular, I used the Jetstream resource at Indiana University through allocation TG-SES170018.

CHAPTER 2: EMPIRICAL FOUNDATIONS

School choice and inequality

Impact on racial and socioeconomic divisions is particularly concerning for charter schools, which were originally conceived (Shanker 1988) and are continually justified (e.g., Pendergrass & Kern 2017) as a bottom-up organizational means to equalize educational opportunity for ethnic minorities and the poor. Thus, concerns for inequality have driven research into charter school practices and outcomes, examining whether they select high-achieving students (e.g., Lacireno-Paquet et al. 2002), negatively affect resources and outcomes in nearby school districts (e.g., Preston et al. 2012), restrict access based on information and social connections (e.g., Yettick 2016), increase educational instability for the under-privileged (e.g., Paino et al. 2017), or sort disadvantaged students into less effective programs (e.g., Golann 2015).

Indeed, the question of whether charter schools have achieved their mission of effective and innovative education is especially important for the poor students and students of color increasingly enrolling in them (Berends 2015; Wang, Rathbun, and Musu 2019). Indeed, compared to traditional public schools, in 2015-16 a higher proportion of charter school students was African American (26.6% vs. 14.6%) and Hispanic (31.7% vs. 26.1%), while a lower proportion of charter school students was white (33.3% vs. 49.4%; author's calculations based on National Center for Education Statistics 2019).

In principle, publicly funded charter schools accept strict accountability to content and performance standards enforced by high-stakes assessments (e.g., Ladd 1996) in exchange (e.g., Paino et al. 2014, 2017) for market-like, innovation-boosting autonomy from many regulations (around labor, content, finances, etc.) facing traditional public schools. Thus, in return for loose regulation, charter schools face strict accountability for achievement gains, financial management, and meeting community expectations (Paino et al. 2014). According to school-choice advocates (e.g., Chubb and Moe 1990), the regulatory independence of schools of choice sidesteps the pedagogical and organizational constraints on traditional public schools, thereby improving student outcomes by promoting organizational competition and a range of parental options. Thus, performance and innovation are the twin objectives of charter schools; to survive, they must balance these goals.

However, studies show that charter schools' academic performance is heterogeneous and not consistently superior to that of comparable conventional public schools: depending on region and locale, some charter schools do better than traditional public schools, some do worse, and some do the same (Berends 2015; Wang et al. 2019). Such mixed findings on charter school performance are notable given that most studies (including the highly cited CREDO reports: Raymond et al. 2013; Woodworth et al. 2017) ignore student selection effects into and by charter schools, which bias achievement measures in favor of charter schools. On the demand side, the decision to apply and persevere in charter schools depends on access to transportation, information about educational options and resources (e.g., language support or special education services), and parents' enthusiasm, spare time, and networks (e.g., Frankenberg 2011; Fuller, Arcidiacono, and Kearns 2020). On the supply side, charter schools recruit and retain students selectively through intimidation of applicants, expulsion (especially after receiving state funds but before standardized test time, i.e. in October), insisting they lack the resources to meet student needs, and denying enrollment (especially to low performers or those with special needs; Adamson and Darling-Hammond 2016). These patterns produce a "flipped choice" (Adamson and Darling-Hammond 2016:155) system wherein charter schools compete for high-quality applicants, rather than vice versa—both inverting the pro-charter rhetoric and relegating low performers to traditional public schools.

In addition, both charter and local public schools share the same set of "innovative" practices at the classroom (e.g., individualized instruction, cooperative learning) and organizational levels (e.g., small class sizes, teacher merit pay; Lubienski 2003), and the same characteristics that raise achievement in charters are also effective for their local counterparts (e.g., increased instructional time, high academic and behavioral expectations, teacher coaching, use of data; Berends 2015; Furgeson et al. 2012; Gleason 2017). Furthermore, when compared to traditional public schools in their vicinity, charters are generally not the sole adopters of administrative innovations in the areas of academic support services (e.g., after-school tutoring), staffing policies (e.g., merit pay), organizational structures (e.g., block scheduling), and governance practices (teacher/parent influence on hiring; Preston et al. 2012). In sum, neither performance nor discrete innovations in pedagogy or organization appear to differentiate charter schools from traditional public schools.

Charter schools and segregation. As schools of choice, charter schools are especially vulnerable to parents' self-sorting by race and class—a key social mechanism of segregation. This is well-evidenced by studies documenting high-status groups like whites escaping to schools of choice to avoid integration with low-status groups like people of color (e.g., Renzulli and Evans 2005; Saporito 2003). This has worked against integration efforts for decades (for a review, see Reardon and Owens 2014)—as well as people of color opting for schools where they are over-represented (e.g., Fabricant and Fine 2012; Frankenberg et al. 2017). Also evident is parents' tendency to select schools with peers similar to their own children in race and class (e.g., Holme 2002)—a tendency reinforced by students' preferences for friendships with similar others (e.g., Shrum et al. 1988).

As a result of these trends, charter schools tend to have student bodies that are more homogeneous by race and class than traditional public schools (Malkus 2016; Monarrez, Kisida, and Chingos 2019b). In 2014-15, 17% of charter schools—compared with 4.5% of traditional public schools—were "racially isolated" (had enrollments that were at least 99% students of color; Moreno 2017). Racial segregation also intersects with class segregation through school choice: 65% of charter schools in 2015-16 were majority students of color, 49% of which had a majority of poor students (receiving free or reduced-price lunch); for traditional public schools, these respective figures are 43% and 34% (Vasquez Heilig, Brewer, and Williams 2019).

Charter advocates counter that charter schools tend to be located in urban areas with concentrated communities of color (e.g., Adamson and Galloway 2019), and thus that segregation in charter schools is largely driven by neighborhood segregation and local geography (Ritter et al. 2010). However, "apples to apples" comparisons of charter schools to traditional public schools located nearby does not erase the trend of their greater segregation by race and class (e.g., Heilig et al. 2016; Lubienski, Gulosino, and Weitzel 2009). In rebuttal, some point to well-funded efforts to create intentionally "diverse-by-design" charter schools—but even advocates can identify only 125 cases (or 2%) of such charter schools (Potter and Quick 2018). Finally, supporters attempt to justify greater segregation among charter schools as deliberate efforts to create culturally coherent, grassroots schools uniquely responsive to their self-segregating clientele (e.g., McShane and Hatfield 2015; Wilson 2016)—or even as sacrosanct expressions of the "civil right" of school choice (Heilig 2013; Scott 2013). In other words, unregulated choice is a higher priority for charter school advocates than are integrated schools (Scott 2018), unlike those traditional school districts that push for desegregation through equity-based policies and financial incentives (e.g., Orfield and Ee 2014).

Such segregation poses several risks: It may deepen inequalities by excluding disadvantaged students from the resources and benefits of integrated schools (e.g., Frankenberg et al. 2019; Hanushek, Kain, and Rivkin 2009), including access to experienced teachers and advanced course work (NAACP 2017); isolate youths from civic engagement opportunities (e.g., Levinson 2012); and lead to ethnic fragmentation, undermining both the "common schools" ideal and democracy itself (e.g., Asante and Ravitch 1991). Thus, it is especially important to understand the mechanisms of educational segregation by race and class in charter schools, the leading edge in a growing trend (e.g., Owens, Reardon, and Jencks 2016; Reardon and Owens 2014). Indeed, though some use selective data to argue that school segregation has grown little since 1998 (Monarrez, Kisida, and Chingos 2019a), careful empirical work shows that segregation both by race (e.g., Frankenberg et al. 2010; Vasquez Heilig et al. 2019) and by class (e.g., Frankenberg et al. 2019; Reardon and Owens 2014) have steadily increased since 1990.

Inequality and parenting styles. Influential psychological research (Baumrind 1966, 1971) documents different parenting styles—which range from affirming ("permissive") to controlling ("authoritarian"), but ideally moderate these extremes with rationality ("authoritative")—and their long-term consequences on confidence, social skills, and academic success. Advancing this line of inquiry, sociologists have discovered two primary, class-distinct parenting styles (Lareau 2000, 2011): the middle-class approach, "concerted cultivation", is characterized by development of individual talents and rich vocabulary, packed schedules, and parental intervention in schooling. In contrast, the poor and working-class approach, "natural growth", is characterized by emphasis on meeting basic needs, strict discipline, parent/child separation, social free time, and parental deference to school personnel. These contrasting logics of parenting thus bestow distinct cultural skills: for concerted cultivation, skill negotiating with authority figures (teachers and parents), individual initiative, and a "sense of entitlement" in asserting one's needs; and for natural growth, deference to authority, self-management, and a "sense of constraint" in holding back one's opinions.

Children raised in the middle-class, concerted cultivation style are advantaged in turning interactions to their interests, winning accommodation from authorities, and navigating institutions—such as getting extra help from teachers (Calarco 2011) or custom treatment by doctors (Lareau 2011). Moreover, educational institutions respond differently to students based on their class backgrounds, socializing in middle-class students qualities of self-direction, creativity, and assertiveness while instilling in working-class students dependability, obedience, and submission to authority (Anyon 1980; Bowles and Gintis 1976; Kohn 1969). This is evidence not of the superior quality of concerted cultivation, but that it provides a means for privileged parents to secure their children's futures by developing their "cultural capital"—those habits, skills, and knowledge that impart cultural and educational distinction through alignment with dominant class tastes, styles, and institutions (Bourdieu 1977, 1984; Lareau 2011).

In sum, my theoretical account suggests that parents' self-sorting by race and class into charter schools reflects their attraction to particular charter school identities (e.g., Malkus and Hatfield 2017) and racially hued perceptions of school quality (e.g., Reay et al. 2008; Roda and Wells 2013), which in turn are driven by class-specific values (Bowles and Gintis 1976; Kohn 1969) and parenting styles—specifically, "concerted cultivation" and "natural growth" (Lareau

2000, 2011). These values and styles are an under-explored causal factor of charter school segregation, and no prior research has analyzed how these correspond with charter schools' identities. This study is an important first step in documenting this theory, by providing meso-level, organizational evidence of the linkage between educational ideology and sociodemographic factors.

Hypothesis

IBL shares with the white, middle-class "concerted cultivation" style (e.g., Lareau 2011) and middle-class educational values (e.g., Kohn 1969) a focus on individual skills and capacities, especially critical thinking; questioning of authority and outside knowledge; and strategic—rather than purely directive—adult guidance of child-centered activities. As such, I predict that:

> Hypothesis: Emphasis on IBL is negatively associated with charter school enrollments of (a) low-income students or (b) students of color.

CHAPTER 3: THEORETICAL FOUNDATIONS

As an increasingly popular fixture of today's educational reform agenda, publicly funded charter schools accept (supposedly) strict accountability to content and performance standards enforced by high-stakes assessments (Ladd 1996; Linn 2000) in exchange (e.g., Paino et al. 2014, 2017) for market-like, innovation-boosting autonomy relative to conventional public schools (from labor regulations, content restrictions, financial oversight, etc.). While overall, charters appear little different from traditional schools in terms of student outcomes (Berends 2015; Raymond et al. 2013) and practices (Lubienski 2003; Preston et al. 2012), charters energetically respond to the expectation for innovation through a variety of pedagogical orientations and philosophies—defining features of their identities (cultural conceptions projected to external audiences, namely parents, students, and state oversight agencies), represented in their mission statements (King et al. 2011; Renzulli et al. 2015).

However, only a few studies (King et al. 2011; Renzulli et al. 2015) have examined identity dynamics in the charter school sector, despite widespread contention over the educational and social effects of contemporary, accountability-based educational reform (e.g., Carnoy et al. 2005; Kamenetz 2015; Mehta 2013; Ravitch 2010). Are charter schools trapped in the same homogenizing "iron cage" (Weber 1905) of rationalism and accountability as the more highly regulated, traditional public schools, or do they succeed in their stated mission of—in the words of one typical state charter law (Lubienski 2003)—"[providing] innovative learning opportunities and creative educational approaches to improve [students'] education" (M. G. Assembly 2003)? In other words, are charter schools as diverse as proponents claim (e.g., McShane and Hatfield 2015; Petrilli 2012), or have the countervailing pressures of isomorphism under the accountability regime (rooted in state standards, high-stakes exams, and incentives; Ladd 1996; Linn 2000) won out? In beginning to address these questions, my quantitative results (see below) provide a first, culturally nuanced glimpse of charter school identities embedded in social contexts.

Indeed, amidst the dearth of empirical research at the nexus of organizations and education (Brint 2013; Renzulli 2014), particularly lacking is research quantitatively examining the organizational heterogeneity, political embeddedness, and local responsiveness of charter schools (Arum 2000). Two competing perspectives have emerged to address these issues. First, in tune with broader neo-institutional analyses of educational organizations (e.g., Arum 2000; Meyer and Rowan 1977; Ramirez and Meyer 2012), organizational institutionalists argue that charter schools are constrained by political and institutional rules and norms, forcing conformity and preventing them from developing situated responses to community needs (Huerta 2009; Zhang and Yang 2008). In contrast, class-stratification scholars argue that charter schools, in order to attain legitimacy and secure scarce resources, differentiate to fill ethnic- and class-specific niches within a developing organizational field (e.g., Fuller 2009; Lauen et al. 2015; Miron and Urschel 2010). Thus, institutionalists view charter schools through an open-systems "bureaucratic" frame of tight, rule-based coordination (i.e., charter schools are more similar than different), while stratification scholars view them through a competitive, technically driven "decentralized" frame capable of greater locally situated responsiveness (i.e., charter schools are more different than similar; Huerta and Zuckerman 2009).

In my empirical analysis (see later chapters), I take the class stratification perspective, analyzing how charter school identities reflect their sociodemographic niches. In this chapter, I discuss identity dynamics among charter schools in their institutional context—in particular, the dominant logic of school accountability and its consequences for educational ideology. I also draw on institutional theory to set the stage for my analysis of organizational mission statements (e.g., King et al. 2011; Morphew and Hartley 2006), which crystallize both broad-based normative appeals (*ritual*; Meyer and Rowan 1977) and identity elements responsive to target sociodemographic audiences and their beliefs and values (*ideology*; Pettigrew 1979).

Institutional analysis: Theories of organizations and politics

Early "new institutionalism" (DiMaggio and Powell 1991) was a break with the past, inspired by influential theories on the "social construction of reality" (Berger and Luckmann 1966)—especially the concepts of *institutionalization* (the process by which beliefs and practices become objectified, habitual, and seen as external, persistent, and immutable) and *legitimation* (the process of explaining and justifying institutions to grant them not only cognitive sensibility but also social uprightness or dignity). In contrast to the focus in the "old institutionalism" (DiMaggio and Powell 1991) on managing local dependencies, intra-organizational politics, and formal-informal conflicts (e.g., Gouldner 1954; Merton 1940; Selznick 1949, 1957), the founding statement of neo-institutional theory (Meyer and Rowan 1977) stresses the point that formal structure itself is a mythical modern image: organizations maintain legitimacy by adopting institutionalized formal structures, regardless of their applicability to local contexts. It is through such symbolic conformity to rationalized rules that formalized organizations secure access to resources and survive (Meyer and Rowan 1977) in a world system that privileges Enlightenment models of means-ends rationality (Meyer et al. 1997). While symbolic formal structures may conflict with efficiency (DiMaggio and Powell 1983; Meyer and Rowan 1977), organizations resolve this contradiction by decoupling their symbolic structure from their functional, technical core (e.g., by making goals ambiguous and not enforcing formal requirements; Chaves 1996; Meyer and Rowan 1977; Weick 1976). Decoupling is concealed, however, by lack of evaluation and a logic of confidence and good faith (Meyer and Rowan 1977). Other formative work in early neo-institutionalism details the coercive, normative, and mimetic processes of isomorphism (DiMaggio and Powell 1983) and the similar regulatory, normative, and cognitive pillars of the legitimate institutional order (Scott 2001).

For early neo-institutionalism, institutionalized formal structures constitute the firm, leading to isomorphism by unreflective, ritual conformity (Meyer and Rowan 1977) as well as by market competition and shared field positions—suggesting that organizational actors trust only those legitimate courses of action documented by experts or those with official credentials (Zucker 1977), adopted by successful others, and encouraged by powerful interest groups (DiMaggio and Powell 1983). Indeed, following work on the bounded rationality of decision-making (March and Simon 1958; Simon 1946), neo-institutionalism holds that formal structure is cognitively bounded: actors can see no alternative to institutionalized formal structures, and when they can, non-institutionalized alternatives are too upstart, unwise, or illegitimate to seriously consider.

Subsequent elaborations of neo-institutionalism, however, have found that the extent of decoupling—meaning ritual conformity, or the failure to implement newly adopted formal structures—varies with the conditions of adoption, such as whether the formal element was adopted early in its institutionalization (more likely to be motivated by efficiency concerns) or late (more likely to be mimetic) (Tolbert and Zucker 1983; Westphal and Zajac 1994; Westphal, Gulati, and Shortell 1997; see also DiMaggio and Powell 1983), and whether the organization is in an "institutional" sector (e.g., schools) or a "technical" sector (e.g., manufacturing) (Meyer and Rowan 1977; Scott and Meyer 1983). Others have noted that seemingly superficial structural changes can lead to substantial intraorganizational changes, such as by reshaping role structures or the gradual adoption of internalized interests (e.g., environmentalist offices may diffuse sustainable values throughout the organization) (Edelman 1992; Scott 2001; see also Selznick 1949). Thus, even where the strategy of ritualized conformity is followed, it may not be successful in divorcing symbolic and technical structures.

Institutionalism and education. Classic neo-institutional theory considers education a prototypical "institutional" social sector (Scott and Meyer 1983) pervaded by decoupling of formal structure from technical practice (Weick 1976), ritual conformity via ceremonial adoption of formal structures to secure legitimacy (Meyer and Rowan 1977), and structural convergence through coercive, mimetic, and normative mechanisms (DiMaggio and Powell 1983). This view emphasizes the symbolic and formal aspects of schooling, which rationalize education (rendering it measurable, objective, and calculated; Weber 1905) but disconnect from the messiness of practical life. In short: "Education is a certified teacher teaching a standardized curricular topic to a registered student in an accredited school" (Meyer and Rowan 1978:219).

However, recent applications of institutional theory (e.g., Coburn 2004; Hallett 2010; Paino 2018) suggest a "recoupling" between the firm hand of state control (see below) and the moldable clay of structure. Although research suggests the chance of school closure is influenced by academic performance only after accountability to financial norms and community expectations (Paino et al. 2014), coercive isomorphism (driven by compliance with resource providers; DiMaggio and Powell 1983) is evident when schools develop methods of instruction aligned with testing requirements and traditionalist models of knowledge (e.g., drills, stressing writing conventions and vocabulary; see below and Watanabe 2007). Moreover, faced with uncertainty, efforts to ensure survival by adhering to legitimated identities lead charter schools into recognizable identity clusters (King et al. 2011; Renzulli et al. 2015), demonstrating mimetic isomorphism (driven by mimicking prominent exemplars; DiMaggio and Powell 1983).

Finally, the growing prevalence of charter management organizations (CMOs) could lead to normative isomorphism (driven by shared administrative or professional standards; DiMaggio and Powell 1983). Due to heavy philanthropic funding (e.g., from the Gates Foundation), CMOs have expanded to account for 32% of charter schools in 2017 (Woodworth et al. 2017), with some estimates as high as 45% (National Alliance for Public Charter Schools 2018), and some boast impressive networks of schools.[3] These schools' freedom from bureaucratic controls explains the growth of large, for-profit CMOs such as Charter Schools USA (serving 48,407 students as of 2017) and the now-defunct White Hat Management (served 14,179 students in 2017; Woodworth et al. 2017). Research has emphasized relationships with stakeholders and internal organization in explaining CMOs' outcomes (Woodworth et al. 2017) and pathways to scaling up (Farrell, Wohlstetter, and Smith 2012).

---

[3] For instance, the KIPP network includes 242 schools as of April 21st, 2020 (see https://www.kipp.org/).

CMOs usually directly control operations and hold charters (like school districts do) for at least three charter schools, intended to provide economies of scale and curricular coherence—while sometimes harnessing their autonomy to extract profits from educational funding (Farrell et al. 2012; Woodworth et al. 2017). CMOs may be strategic to protect their schools from threats: for instance, they could prevent charter revocation by improving their schools' academic results (Woodworth et al. 2017) and ensure reauthorization through effective fiscal management (Paino et al. 2014). Indeed, CMOs are an emergent organizational form with increasing political legitimacy, and as such their schools likely enjoy legal and reputational advantages for survival (Edelman 1992). Responsible for such benefits are the teams of executives, content providers, and bureaucrats running CMOs—just as they do other large organizations. And given the well-known tendencies of bureaucrats to regard established professional norms and rules as sacrosanct (DiMaggio and Powell 1983; Merton 1940; Weber 1968), to hold notions of effectiveness or "best practices" in line with managerial fads (Abrahamson 1996), and so on, their influence could limit innovation and promote uniformity in their member schools and beyond (e.g., Goodman 2013; Hassel and Peterson 2006).

To further elaborate the causes for "recoupling" between classroom practice and political-organizational discourse (Coburn 2004; Spillane and Burch 2006), I now discuss the ascendant logic of accountability and its consequences.

The rise of accountability. In modern capitalist societies, formal education is a key determinant of life chances and a dividing line between social strata (Becker 1962). Matching the social and economic importance of education has been its constant political cachet: dissatisfaction among working classes and lower social strata constantly antiquated (rendered "uneducated") under advanced capitalism has led to waves of educational reform to crest within and across nations (Adamson, Astrand, and Darling-Hammond 2016) Such reforms have included Progressive and equity-based programs, but the most enduring have been those adopting the rationales of neoliberalism (Mehta 2013). Surging through such efforts has been the organizing logic of accountability.

Accountability is the notion that schools and school districts should be held to account for meeting organizational standards, primarily academic performance (Paino et al. 2014). It is thus a mechanism for rationalization through which education becomes measurable, objective, and calculated (Mehta 2013; Weber 1905). By tying key provisions of the welfare state (public education) to "objective" metrics for assessing quality (academic outcomes), results-driven accountability provides the core neoliberal logic behind decentralizing educational provision, decreasing public investment, and reliance on private and for-profit solutions. Indeed, this organizing logic dovetails with standards-based educational reforms and school choice, core principles in the international political-economic effort toward educational privatization (Adamson et al. 2016).

These developments in education are part of a larger "accountability culture" rooted in neo-liberal ideals of economic efficiency (Strathern 2000), which have also fueled such market-based reforms as charter schools and voucher programs. Indeed, the dominance of the accountability logic was a precondition for the educational reform of charter schools starting in the early 1990s. Advocates of market reforms argue that parental choice, school competition and autonomy, and market control are prerequisites for improving school performance (Chubb and Moe 1988, 1990). In their view, bureaucratic structures hinder school responsiveness: reforms depending on democratic hierarchical control impose on schools values agreed upon through state or national discourse rather than in response to local constituents and interest groups, greatly restricting the degree to which public schools can effectively serve their communities (e.g., by limiting parents' voice in principal selection or budgetary planning). Privatized educational arrangements are thus a more efficient answer to the question of how best to prepare democratic citizens, preventing unnecessary government involvement in school operations as well as addressing systemic inequities in education access (Friedman 1955).

Through the lens of accountability, then, government waste is minimized and market efficiency is maximized when consumers (parents and students) choose those schools most effective at providing quality (higher standardized test scores), leaving poorly performing schools fewer resources and eventual failure (Chubb and Moe 1990; Paino et al. 2014). This broadly appealing logic connects not only provinces within nations, but also across nations due to widespread acceptance of international tests (PISA, TIMSS). It has led to a panoply of reforms that are publicly funded but free of local authorities' control, including charter schools in the U.S. and academies in the U.K.; school vouchers in Sweden, Chile, and Washington, D.C.; and for-profit school operators in the U.S. and elsewhere. Absent coupled reforms in educational funding and social welfare, however, such efforts have little payoff beyond undermining the welfare state—deprofessionalizing teachers, homogenizing schools, and failing to deliver gains and reduce gaps in achievement (Adamson et al. 2016; Johnson 2019; Mehta 2013).

The current political cycle of educational standards and accountability in the U.S. began with *A Nation at Risk*, a report sounding the alarms over national and economic security given flagging school performance (Mehta 2013). This report inspired a new educational accountability movement, which was taken up by state-level actors in the (especially late) 1980s and 1990s; by the time the federal No Child Left Behind (NCLB) act was enacted in 2002, policies rooted in standards and accountability had already passed in 49 states. Such systems are defined by rewards (financial and/or reputational) and sanctions, which range from student transfer options to reconstitution (replacement of all school staff) and even school closure. Although its goal of improving student outcomes is popular beyond debate, the unintended local consequences of accountability include teaching to the test (Watanabe 2007) and prioritizing profits or fiscal benefits over academics (Paino et al. 2014). Indeed, demonstrating achievement and standards compliance has become so urgent as to lead schools to cheat on standardized tests by altering scores and sending poor performers home before test time (Adamson, Cook-Harvey, and Darling-Hammond 2015; Blinder 2015; Fernandez 2012).

The accountability wave continues under Obama's Race to the Top legislation (2009) and the Common Core state standards (first released 2010 and now in 44 states). Moreover, although the recently-implemented (Green 2017), Obama-era Every Student Succeeds Act (ESSA) requires that states choose a "student quality or student success" indicator, the bulk (four out of five) of the school performance indicators are individual-level academic measures: academic achievement, academic progress, English proficiency, and high school graduation rates. In fact, every state has chosen as their student quality/success indicator another easily observed, individualistic measure: either chronic absenteeism/attendance or some predictor of college readiness like SAT/ACT scores or access to Advanced Placement courses (Jordan and Miller 2017; Sawchuk 2017). Furthermore, though politics may retune the details of school and teacher accountability—for instance, as the ESSA shifts regulatory control from the federal government to the states, or as Congress overturns clarifying accountability guidelines as they did in early 2017 (Goldstein 2017)—accountability-based policies continue to transform U.S. education.

And yet the accountability logic is far from uncontested. They emerge through the mobilization of state officials and business and parent organizations—over the continuing objections of teachers' unions and local school authorities (Evers 2001). In particular, teachers tend to oppose accountability measures in defense of their professional autonomy against oversight and standardization (Drudy 2008; Johnson 1972; Guarino, Santibañez, and Daley 2006). But today's protests by the professionally weak and mostly feminized K-12 teaching force are largely inconsequential politically—just as in previous reform eras (Mehta 2013, 2014).

Moreover, the triumph of the accountability logic meant the subordination of norms and values internal to the education sphere to an external, rationalizing logic based on quantifiable measures (Mehta 2013:10–38). In this effort, reformers transplanted into education the economic logic of inputs and outputs, profits and losses, and successes and failures from the high-status business domain (Evers 2001; Mehta 2013). This shift in logics was a constant across states' routes to accountability-based educational policies, resulting in remarkably similar, punitive practices despite divergent motivations and contexts (Mehta 2013; Watanabe 2007). Thus, school choice policies and their resulting organizational dynamics emerged as explicit responses to the pressures of accountability.

Accountability and educational ideology. What sorts of identities do charter schools proclaim? Based on influential arguments for progressive education (Dewey 1938), popular discourse around educational philosophies (e.g., Kohn, Meier, and Loveless 2006; Meier 2013b), analyses of charter school mission statements (e.g., King et al. 2011; McShane and Hatfield 2015), and studies of how teachers and schools interact with accountability programs (e.g., Evers 2001; Hallett 2010; Watanabe 2007), I distinguish between two foundational ideologies that apply to charter schools: traditionalism and progressivism. In broad outline (Oakes et al. 2013; Watanabe 2007), traditionalism calls for disciplined, scientifically organized, and often rote development of the knowledge and skills needed for student success in college and careers, with primacy given to "core subjects" in the Western tradition (or "common cultural canon") including mathematics, literature, and grammar (e.g., Harris 1906; Hirsch 1996). In contrast, progressivism advocates a more "well-rounded", "child-centered" approach meant to instill social and emotional competence, independent thinking, and civic participation through active and self-directed learning (e.g., Dewey 1897). These conflicting paradigms reach broadly into educational discourse, each visibly manifesting through discrete philosophies and curricula. For example, traditionalism includes No Excuses, back-to-basics, and STEM-based (Science, Technology, Engineering, and Math) schools, while progressivism includes multiculturalism (e.g., bilingual education), project-based learning, and Montessori and Waldorf schools.

How does educational policy shape how teachers implement these ideologies? In addition to consistent self-reports that teachers emphasize tested subjects (e.g., Fusarelli 2004), research grounded in classroom observation and teacher interviews (Watanabe 2007) finds that state standards considerably influence teaching practices, narrowing teachers' pedagogical focus to those standards most clearly reflected in the state's standardized tests (e.g., structure and elaboration in writing, vocabulary) and otherwise displacing even those teaching goals both advocated by teachers and formally espoused in the state standards themselves (e.g., collaborative learning, appreciation of literature). Teachers do have some agency in this process, however: as "street-level bureaucrats" committed to direct service despite the constraints of the public bureaucracy that employs them (Lipsky 1980), teachers navigate standards on the basis of personal and professional beliefs, preferences, identities, and knowledge (Coburn 2004; Goldstein 2008; Sloan 2006) and develop educational and political strategies responsive to organizational context (Hallett 2010).

However, the institutional order of accountability is built on and reinforces an approach to education that discourages such teacher innovation: the pedagogical philosophy of traditionalism. In the view of traditionalism, the goal of education is for students to develop (regardless of how) an essential core of finite, measurable knowledge and skills required for success in today's world (Hirsch 1996; Oakes et al. 2013; Watanabe 2007). In this view, resolving the fundamental issue of low standards and expectations by professionally weak teachers (Kornhaber 2004; McNeil 2000; Meier 2002)—a key source of the perpetuation of educational and social inequalities in class and race—requires the heavy hand of standardization, incentives, and high-stakes tests.

The ideology of traditionalism is aligned with the logic of accountability. They are mutually reinforcing in a coevolutionary pattern (Haveman and Rao 1997): accountability facilitates traditionalist "instructional program coherence" (Newmann et al. 2001), while traditionalism gives accountability "teeth" in the learning environment. While traditionalism calls for dedication to hard skills like writing conventions and vocabulary, accountability incorporates performance metrics into a system of incentives for schools, teachers, and students (Linn 2000). While traditionalism demands disciplined, orderly focus on core subjects (i.e., math and reading), accountability provides an infrastructure of standards and assessments that narrows pedagogical attention to the subjects emphasized on standardized tests (Watanabe 2007).

Favored by its congruence with the logic of accountability, traditionalism has become a legitimate and exceedingly common basis for education. For instance, the dictates of traditionalism are clearly reflected in "No Excuses" schools, exemplified by the KIPP group of charter schools (Angrist, Pathak, and Walters 2013; Meier 2013a). The No Excuses model targets disadvantaged populations (especially racial minorities) and emphasize high expectations for student discipline (often including uniforms and "cold-calling" students for answers—rather than waiting for student volunteers) and academic performance, data-driven instruction, increased instructional time, traditional reading and math skills, feedback and selective hiring for teachers, and a rigorous focus on building human capital to bring college into reach (Angrist et al. 2013; Carter 2000; Fryer 2011; Thernstrom and Thernstrom 2004). No Excuses schools are possibly the most popular specialist model in the whole country (in terms of both school and student counts; McShane and Hatfield 2015).

In addition, the logic of accountability depends on and reifies core values of American education, specifically those that have been taught by public schools for decades to bridge youth from the world of family to that of work (Dreeben 1968; compare with Bowles and Gintis 1976 and Durkheim 1961). These are independence (solitary effort on academic tasks, especially formal tests), achievement (students accept and learn to identify with formal evaluation of their academic efforts), universalism (expecting objective and "fair" treatment based on the general role of "student", rather than personal characteristics), and specificity (accountability for performance on discrete tasks, rather than global, holistic evaluation) (compare with the more community-oriented, egalitarian, and intimate "free schools"; Swidler 1979). Such bedrock values clearly support student discipline in test-taking situations and tolerance for academic challenges. Moreover, the hegemonic model of educational socialization, called "concerted cultivation" (Lareau 2011), is in harmony with accountability-era values of efficiency, competitive advantage, and academic achievement.

Yet the individualistic orientation of concerted cultivation also resonates with the progressive educational model (Dewey 1897; Meier 1995; Oakes et al. 2013), the basic counter-tendency to traditionalism. Like concerted cultivation and its emphasis on unique interests, progressivism values education that is not standardized but student-centered and participatory, not formally disciplined but caring and non-hierarchical, and not focused on drilling basic skills but exploratory and based in social interaction. While traditionalism relates to several basic values of educational socialization (independence, achievement, specificity, and universalism; Dreeben 1968) and may reflect workplace discipline to come (Bowles and Gintis 1976), the educational philosophy of progressivism may better reflect another core value of U.S. education and the "cult of the individual" (Durkheim 1912): uniqueness. This value, which holds the "self" as the context of learning and the source of new capacities, is taught in the U.S. as early as preschool (Tobin, Wu, and Davidson 1991) through emphasis on self-reliance, self-expression, individual talent, and free choice. In sum, the "virtuous self" (Frye 2012) shaped by educational socialization in the age of accountability is self-sufficient, academically skillful, and task-focused (Dreeben 1968); possesses a strong ethic of fairness and highly developed, unique individual talents and interests (Tobin et al. 1991); and is confident in confronting authority and effective in managing time commitments (Lareau 2011).

Many teachers hold progressive views (e.g., Evers 2001; Watanabe 2007), and what limited studies there are (Malkus and Hatfield 2017; McShane and Hatfield 2015) suggest that schools with a "progressive" (i.e., "project-based" or "child-centered") mission are about just as numerous as are No Excuses schools. The roughly equal presence of these two general educational approaches may seem surprising, given that traditionalist educational models are aligned with the logic of school accountability that dominates educational policy (e.g., Evers 2001; Watanabe 2007). Amidst an apparently adversarial political environment, the sustained appeal of progressivism likely emerges from its patchwork of educational ideologies, which are relatively distinctive from each other (rather than a standardized "big tent") and more particular to market niches or sociodemographic communities (potential "customers"). For these reasons, in some statistical models this project uses a single progressive ideology (IBL) to test whether charter school identities distinguish to what niches charter schools advertise themselves.

Organizational identity

A core insight in the organizational identity literature is that organizational survival depends on effectively balancing two identity constraints: appearing unique while also fitting in. On one hand, organizations must be internally coherent, developing expectations for a "central, enduring, and distinctive" identity (e.g., vision and values; Whetten 2006) and remaining true to those expectations or historical "imprints" (Johnson 2007; Stinchcombe 1965) in the eyes of internal stakeholders (e.g., employees; Hannan et al. 2006). On the other hand, organizations must express this distinctiveness appropriately relative to acknowledged rivals in the field (Porac et al. 1995) and the institutional expectations of external stakeholders (e.g., the state or consumers; DiMaggio and Powell 1983) in order to fit an established niche and claim legitimacy (e.g., as a specialist in an established genre; Hsu 2006; Swaminathan 2001; Zuckerman 1999). This general process of balancing group assimilation and differentiation is encapsulated in optimal distinctiveness theory, which holds that individuals and groups seek balance somewhere between uniqueness and homogeneity—or being "just different enough" from both in-groups and out-groups (Brewer 1991)—and has been demonstrated at the individual, collective, and organizational levels (for a review, see Leonardelli, Pickett, and Brewer 2010).

The charter school sector is an ideal setting to analyze identity dynamics because of its pseudo-market orientation, which foregrounds innovation and uniqueness while subtly encouraging isomorphism through coercive (state performance standards), mimetic (legitimated identity clusters), and normative (organizational-professional) pressures (DiMaggio and Powell 1983; see above). Thus, charter schools face conflicting institutional logics (symbolic systems that order reality and direct action; Friedland and Alford 1991). On one hand, the pressure for performance generated by the accountability logic (Hallett 2010; Mehta 2013) demands disciplined focus on essential, testable knowledge in core subjects like math and language arts. On the other, pressure for the innovation expected under the market logic (Fligstein 1990) invites organizational and pedagogical experimentation. See Appendix A for a visual representation of these dual pressures—for uniqueness and community responsiveness, on one hand, and homogeneity through accountability and standardization, on the other.

As discussed above, according to canonical neo-institutional theory, this tension is resolved through ritual conformity (adoption of institutionalized formal structures in order to secure legitimacy) and decoupling (disconnection between symbolic structures and the technical core; Meyer et al. 1997; Meyer and Rowan 1977; Weick 1976). However, recent research suggests such strategies are increasingly difficult to pursue (e.g., Coburn 2004; Hallett 2010; Paino 2018; Watanabe 2007), and as such likely do little to blunt these pressures.

Evidence on charter school identities. Empirical studies do suggest tension among charter schools between conformity with field-level expectations and organizational-cultural innovation. Analysis of charter school mission statements in Arizona from 1996 to 2001 (King et al. 2011) discovered convergence of charter schools over time toward two distinct identity clusters: on one hand, creative arts and learning, which focuses on the arts, creativity, values, recreation, and a specific curriculum (e.g., Montessori); and on the other, family and social programs, which targets at-risk students, emphasizes ethnic identity, and offers special services (e.g., vocational, legal, social). In addition, the odds that a charter school develops the elements of a coherent identity cluster were found to increase with the prevalence of that cluster in the school's district, suggesting localized isomorphism. However, this study also finds that charter schools create innovative identities by combining, on one hand, the qualities associated with membership in the two identity clusters described above; and on the other, "flexible elements" unattached to any particular identity category (e.g., general themes like college-bound and science and technology or basic services like computers and library access; King et al. 2011:560–61). They also show that access to alternative organizational templates (i.e., the presence of magnet schools, which offer innovative programs to attract diverse students) decreases localized isomorphism by introducing heterogeneity into charters' institutional resource base.

Moreover, a broader study of charters' (abbreviated) mission statements in the U.S. from 1992 to 2005 (Renzulli et al. 2015) found that specialized charters (e.g., those with focus on the fine arts, special education, languages, and/or vocational training) became less common over time relative to generalists (those with no clear specialization in curriculum, thematic focus, or target population). This demonstrates assimilation of specialists—more so than their deaths—into an increasingly generalist-dominated charter sector. However, this study also documents specialist differentiation: while generalism increases, those charter schools with specialist missions—largely due to newly opening schools—become more and more unlike each other over time. However, this trend appears to flatten out shortly after 2001, suggesting that strict school accountability under No Child Left Behind limited differentiation of charter school identities (Renzulli et al. 2015:94).

These findings suggest a complicated—and incomplete—narrative. Early in the history of the charter sector, these schools tended to copy each other (within limits) and consolidate identities into recognizable clusters (King et al. 2011). Over time, with the nation-wide establishment of the charter sector and the consequent legitimacy available even to charter schools without any particular claim to innovation (in the form of specialization), it is possible that reliance on identity clusters declined, leading to the proliferation of generalists (as documented in Renzulli et al. 2015). Other research also suggests as many as half of charter schools are generalists (McShane and Hatfield 2015), with mission statements I would classify as purely ritualistic.[4] However, the prevalence of generalism may be over-stated by reliance on the simplified mission statement categorizations of a third-party entity (the Center for Education Reform; Renzulli et al. 2015:88), rather than observing the mission statements explicated by charter school organizations themselves (e.g., as on their websites, the source of mission statements in the present research).

What identity dynamics are at play now that both the charter sector and the era of school accountability (see above) have come into maturity? Has deference to accountability become an all-important ritual in establishing a legitimate identity—or do meaningful differences continue to separate clusters of charter schools, as they once did in some places (e.g., King et al. 2011)? Research thus far leaves these population-level questions unanswered, despite their importance to the sociology of education (in particular, around charter school organization and efficacy; e.g., Carnoy et al. 2005; Chubb and Moe 1988; Paino et al. 2014), institutional analysis, and organizational identity literatures. I begin to address these questions through an analysis of charter school identities, building on research into organizational mission statements—a culturally rich artifact I now briefly describe.

Mission statements: Part ritual, part ideology. In charter schools research, the term "mission statement" has been applied to texts ranging from application (King et al. 2011) or website self-descriptions (McShane and Hatfield 2015) to third-party summaries of schools' educational approaches (Renzulli et al. 2015). Thus, by "mission statements" I refer not only to highly condensed statements of purpose (e.g., "Our school will prepare children for successful futures"), but rather to organizational self-descriptions broadly speaking—including website texts (as in my corpus) including school philosophy, creed, vision, values, and other heterogeneous identity claims.

It is the dual, often-competing requirements for enduring uniqueness, on one hand, and context-dependent conformity, on the other, that organizations actively condense (usually by committee deliberations and often with community input) into expressions of purpose and identity that appeal to multiple audiences: mission statements. Mission statements are expressions of identity at multiple levels: the internal (Albert and Whetten 1985) as well as the institutional, relating themselves to established labels, categories, and reference groups (King et al. 2011; Porac et al. 1995; Rao, Monin, and Durand 2003; Whetten 2006; Zuckerman 1999). Although mission appears a "soft" organizational feature compared with administrative routines,

[4] An example of a highly ritualistic mission statement is: "PUC's mission is to uplift communities through the creation of high quality public charter schools in which students are inspired and prepared to graduate from high school and university and commit to uplift their communities now and forever" (http://www.pucschools.org/tchs/). For longer examples, see Appendix C.

leadership hierarchies, etc., mission statements meaningfully shape organizational life—as well as "hard" features like performance. For instance, the practices implemented by the Knowledge is Power Program (KIPP) and other schools identified with the "No Excuses" framework—a defining feature of which is a "high-achieving culture"—generate substantial test score improvements (Angrist et al. 2013; Tuttle et al. 2010). Moreover, state accountability plans under the Every Student Succeeds Act are permitted to measure school quality partly through campus climate, providing a foothold for such "soft" features in the legal framework of accountability (e.g., Sawchuk 2017).

In recognition of the layered appeals and levels of meaning in organizational mission statements, I reason that mission statements mix two basic ingredients: the ritual and the ideological. "Ritual" elements constitute broad-based normative and symbolic appeals not relevant to the organization's specific history, structure, or context (e.g., use of the term "health" across medical schools, "service" across colleges and universities, or "patient care" across hospitals; Grbic, Hafferty, and Hafferty 2013; Morphew and Hartley 2006; Williams et al. 2005). Indeed, the prevalence of such ritualism has led some observers to view mission statements as providing "vague and vapid goals" (Chait 1979) or "honorable verbiage signifying nothing" (Newsom and Hayes 1991:29). In contrast, "ideological" elements reflect idiosyncratic beliefs, values, and goals that may be charged with meaning by the organization's history, values, or social agenda (e.g., reference to social justice principles inspired by abolitionist founders; Hartley 2003). Intermediate in this ritual/ideology polarity are those institutionally specific signals meant to appeal to shared, important constituencies. For instance, the mission statements of public colleges and universities signal service to state benefactors by using the element of "serves local area", while the mission statements of private universities signal technical superiority to affluent potential customers with terms like "leadership" and "student development" (Morphew and Hartley 2006). Similarly, the empirical goal of the present study is to discover in charter school identities targeted, differentiating appeals to community-specific educational ideologies.

While construction from ritual elements makes mission statements formalized, symbolic artifacts in sync with prevailing notions of rationality but potentially divorced from organizational life (Meyer et al. 1997; Meyer and Rowan 1977), construction from ideological elements can offer the counter-balance of expressing myths and meanings socially constructed within the political, cultural, and symbolic ecology of the organization (Pedersen and Dobbin 2006). It is this mix of elements that makes mission statements an excellent tool for studying organizational identity, for in a single text they hold legitimacy claims (ritual), uniqueness claims (ideology), intermediate elements or community claims (constituency appeals), and the relative proportions of all these. Moreover, given their near universality amongst charter schools (via websites), mission statements allow the study of identity dynamics amongst the entire U.S. charter school population.

CHAPTER 4: RESEARCH METHODS

Previous attempts to classify charter school identities have relied on hand-coding limited samples (e.g., McShane and Hatfield 2015; Renzulli et al. 2015), resulting in several incongruent categorizations—e.g., a set of 13 categories such as No Excuses, international, and arts (McShane and Hatfield 2015) versus a very different set of 11 categories such as values, homeschool, and special education (Renzulli et al. 2015). These varied categorizations imply charter schools are very active in differentiating themselves, resulting in mission statements that are less ritualistic and more ideological than in many of the other settings studied, such as colleges and universities (Morphew and Hartley 2006), medical schools (Grbic et al. 2013), and hospitals (Williams et al. 2005). Prior research (e.g., King et al. 2011; Renzulli et al. 2015) has distinguished four elements that charter school mission statements provide to demonstrate innovation: curriculum (e.g., Montessori or college-oriented), thematic focus (e.g., STEM or marine biology), target population (e.g., gifted or at-risk students), and resources and services (e.g., arts facilities or full-day kindergarten). Other studies have compared charter schools using simple categorical schemes of market "niches" such as district-affiliated vs. not, start-up vs. conversion (Lauen et al. 2015) or critically examined the racial implications of marketing materials from a few large charter management organizations (CMOs)[5] (Hernández 2016). While such studies are illustrative, they significantly reduce the complexity of social contexts, obscuring charter schools' embeddedness in communities of varying demographic characteristics.

Moreover, the scale and depth of previous studies has been limited either by a small number of research sites (e.g., Oakland school district: Jha and Beckman 2017; or Arizona state: King et al. 2011) or the superficiality of the cultural information examined (e.g., by sorting schools into preconceived categories: Malkus and Hatfield 2017; or by relying on third-party summaries of charter school missions: Renzulli et al. 2015). The difficulty of collecting comprehensive, valid data and the sensitivity of measurement to geography and history impede effective, theoretically grounded analysis of charter school identities. Building on groundbreaking advances in computational social science (e.g., Mohr, Wagner-Pacifici, and Breiger 2015; Nelson et al. 2018), I overcome these methodological obstacles by collecting detailed, comprehensive, ecologically valid data on charter schools and their social contexts and applying flexible, reliable text-analytic methods.

Data and measures

Schools' websites appeal to parents and school district authorities, connect staff, and detail instructional design choices (Bryk et al. 2010): the skills and traits it develops, the behaviors it promotes and restrains, its mission and values, its view of the learning process, etc. Such cultural cues seek to deliver what their audiences expect; they are partial artifacts of organizational impression management (Elsbach 2003). As theorized above, websites contain a mixture of niche-specific, ideological claims together with general-appeal, ritualistic language. See Figure

[5] I define a CMO broadly as any private organization that manages two or more charter schools—so long as this organization's website lists all schools it manages, thus encapsulating them within a larger "brand" identity.

4.1 for the most least distinguishing, most ritualistic words in charter school websites. These words are used frequently, rather than in specific contexts; they contrast with ideological terms that differentiate websites by reflecting community-specific educational beliefs and ideals.

[FIGURE 4.1 ABOUT HERE]

While websites are ubiquitous and culturally relevant, however, user experience research shows that most readers scan pages quickly, retaining at most 28% of their text content (Nielsen and Morkes 1997). And the text people do read is not taken for granted; an organization's self-descriptive claims made in "About Us" pages tend to be checked against third-party sites (Kaley and Nielsen 2019)—a pattern exacerbated by the prevalence of review sites (e.g., greatschools.org, schooldigger.com) in the top, most-viewed search engine results. Because I operationalize the concept of organizational identity as website self-descriptions, the above conditions make it less likely that identity (as measured here) influences parents' school choices. Thus, this study amounts to a conservative test of my hypothesis.

Web-crawling. I used web-crawling in Python 3 to gather data on organizational identities from the websites of all 6,872 charter schools open in 2015-16 (National Center for Education Statistics 2019), about 92% of which had websites when crawled in June of 2018 (author's calculations).

My web-crawling workflow had three steps: collect URLs, crawl web pages, and parse and filter text. My code and URL lists for charter schools and CMOs are available online.[6]

First, given that no comprehensive, reliable list of charter school URLs exists , I used the Google Places API (free for researchers: https://cloud.google.com/maps-platform/places/) to collect URLs. I specifically searched for the name and full address of each charter school, which tended best to distinguish schools and to return distinct, accurate URLs. My research team and I then cleaned this URL list by hand to identify the domain (i.e., URL) most specific to each school that also contained information on its mission, values, etc.—either on the domain itself (e.g., *www.school.com*) or in its subdomains (e.g., *www.school.com/about-us*).

For web-crawling, I specifically used Scrapy Cluster, a scalable crawling architecture for Python that coordinates multiple web spiders (IST Research Corporation 2017). For each domain, I crawled all subdomains by following page links restricted to the root domain (e.g., *www.school.com* could link one level deeper to *www.school.com/page*, but not outside to *www.faceblast.com/school*), followed all links on those subdomains, and so on to a maximum crawling depth of 10 (most sites were not this deep). I crawled each page only once. Each PDF file (if any) was converted into text and considered a page.

Because charter school websites are diverse and lack a consistent, coherent structure, I used `BeautifulSoup` (Richardson 2007) to clean website text using simple HTML parsing: I removed inline or formatting tags (e.g., `span`, `strong`) and non-visible tags (e.g., `style`, `head`), leaving only visible text as would be viewed in a web browser. For this analysis, to represent each school I joined all its crawled, parsed pages into a single string of text. To preprocess the text for all analyses below, I also lower cased all words and removed punctuation

---

[6] See my code for collecting URLs, crawling web pages, and parsing text—as well as a complete list of charter school and CMO URLs—at http://bit.ly/web_crawl_tools.

and numbers (additional filters described in Structural Topic Model section).[7] To accurately represent charter schools in their organizational contexts, I also web searched to create a current, complete directory of CMOs, their URLs, and the schools they manage.

To create a current, complete directory of CMOs, their URLs, and the schools they manage, I first merged two lists of CMOs: one from a recent report (Woodworth et al. 2017), the other shared (upon request) by the National Alliance for Public Charter Schools (https://www.publiccharters.org/). I then manually cleaned and extended this list through web searches, and finally checked CMO websites to enumerate their schools.

Web-crawling yielded data on 6,300 websites, or 91.7% of all open charter schools. Most of these websites have a significant amount of information: 87.5% include up to 100 web pages, and 88.0% have more than 200 words. However, 7.6% of websites have less than 10 words—a weak information source; I remove these to strengthen my measure of school ideology, reducing the sample to 5,806 schools. I address the possible effects of other outliers (e.g., the 12.5% of schools with more than 100 pages) in the "robustness checks" section below.

To handle other missing data for those 5,806 schools whose websites I successfully captured, I implemented multiple imputation (Rubin 1976) using the `mi` package in Stata 15 (StataCorp 2017a) with 100 imputations. Multiple imputation uses predictive modeling to compute multiple sets of plausible values (here, 100) to replace missing data. Each imputation is then analyzed separately, and their estimates are pooled into a single result—with standard errors reflecting the sampling variability between imputations. Thus, multiple imputation is more efficient and representative than listwise deletion and more precise than single imputation. Accordingly, I dropped only 22 cases missing information on school size, demographics, or grade range served, yielding 5,784 schools in my models.

Variables. My first dependent variable is the degree of emphasis on IBL, measured as the percentage of IBL terms on the school's website (see "Dictionaries" section for detail). My second and third dependent variables are the school's percentage of students of color (black, Hispanic, Native American, Asian, Pacific Islander, or multiracial), and the percentage of students receiving free- or reduced-price lunch (FRPL, a proxy for poverty), at the school level. Each of these also serves as independent variables in some analyses. My other independent variables are school district demographics, specifically the percentage of residents of color and the percentage of families below the poverty level; and school academic performance, measured as proficiency rates on standardized state assessments of reading/language arts and mathematics.

To capture these variables, I match web data to school data in the 2015-16 Public School Universe Survey (PSUS) of the National Center for Educational Statistics (NCES 2019). The PSUS data include ethnicity and FRPL, plus the grade range (as dummy variables: primary, middle, high, and other/ungraded), operating status (used to calculate each school's age), number of students (in hundreds; excludes adult education), urban locale status (whether a school is in a

[7] Because computers excel at narrowly defined tasks rather than human-level cognitive understanding (Geiger et al. 2019), computational text analysis necessarily involves reducing the complexity of texts through preprocessing. As a consequence of this dimensionality reduction and the inherent complexities of language, all text analytic methods are "wrong" as models as language (Grimmer and Stewart 2013:3–4) and also often qualitatively "weird" when interpreted by humans (Shane 2019).

central city of at least 50K residents), and latitude/longitude (which I use to geo-locate charter schools into school districts). I also match to data on school districts in the 2012-16 5-year estimates of the American Community Survey (ACS; U.S. Census Bureau 2018), which includes metrics on ethnicity and poverty. Lastly, I match with the EdFacts 2013-14 and 2014-15 school-level proficiency scores in reading/ language arts and mathematics maintained by the U.S. Department of Education (USDE 2018)—high-quality, comprehensive data widely used in education research (e.g., Reardon, Kalogrides, and Shores 2017).

Table 4.1 provides descriptive statistics on these variables, and univariate distributions for key variables are presented in Figures 4.2 – 4.5. Only 0.06% of words on the average charter school website come from the IBL ideology (results not shown), suggesting a rich diversity of language in the corpus. The average charter school has 48.3% and 40.2% proficiency in reading/language arts and math, respectively. Enrollments of poor students and students of color are similarly common: the average charter school enrolls 55.5% poor students and 65.0% students of color. School districts on average enroll fewer poor students (by these measures) and students of color than do charter schools, with 14.7% poverty and 33.5% people of color.

[TABLE 4.1 ABOUT HERE]

[FIGURES 4.2 – 4.5 ABOUT HERE]

Analytic strategy

I use Structural Topic Modeling in R (Roberts et al. 2014; Roberts, Stewart, and Tingley 2019) to inductively discover latent themes in charter school websites. I then deductively predict schools' emphases on IBL with count-based dictionary methods (Grimmer and Stewart 2013; Jurafsky and Martin 2018) and neural-net word embedding models (Mikolov, Chen, et al. 2013) in Python followed by mixed linear models in Stata.

This workflow represents an application of computational grounded theory (Nelson 2017). I begin with content knowledge (segregation in charter schools) and then detect patterns with unsupervised computational methods (topic modeling with simple linear models: Chapter 4). I interpret these results and select a test case (IBL) to confirm the sociological patterns I've induced. In the next iteration (see Chapter 5), I again apply content knowledge (seed terms), which I refine with unsupervised computational tools (word embeddings). Finally, I confirm the discovered patterns with sophisticated models (mixed linear regressions) and various robustness checks, including testing alternative word lists. Indeed, an emerging trend in the social sciences is a similar workflow that starts with domain knowledge, constructs unsupervised models for pattern detection, and seeks pattern confirmation by applying dictionaries of different sizes—through counts or cosine distances—to specific corpora (e.g., movie reviews or social media posts: Garten et al. 2018; Sivak and Smirnov 2019).

In the remainder of this chapter, I outline my topic modeling, dictionary-based, and word embedding approaches before detailing how I built the IBL dictionary. I then describe the mixed linear regression models.

Mixed-membership topic models. Topic modeling comprises a family of popular, automated methodologies used to classify large amounts of text for theoretically-informed exploration or hypothesis confirmation (DiMaggio, Nag, and Blei 2013; Mohr and Bogdanov 2013). I use topic modeling to inductively analyze variation in word clusters from charter school websites in order to identify "topics" that signal educational ideologies (on using automated text techniques to measure cultural phenomena, see generally Bail 2014). Although the so-called "bag of words" assumption – i.e., that the order of words within a document is unimportant – leads topic modeling to ignore word order (and hence grammar, phrasing, logic, etc.), this technique nonetheless can reveal latent textual patterns obscured by other methods (Blei, Ng, and Jordan 2003).

Websites comprise rich sources of data about organizations, yielding "narratives" that signal collective identities, instrumental information, maps of formal structures and informal networks, and governance systems that comprise particular organizational forms (Powell et al. 2016; see also, Bennett and Segerberg 2013; Ferraro and O'Mahony 2012). The capacity of topic models to identify themes from large bodies of text is especially useful given the heterogeneous character of charter school websites, which target multiple audiences while revealing much about each school. These websites contain information on everything from educational philosophy to staff biographies to scheduled events and meal plans. By capturing multiple, mixed-membership themes that separate content areas within and across websites, I expect topics on educational ideologies to coalesce, comprised of word clusters largely exclusive of other topics. Although such information as dress code and disciplinary policies are valuable to some stakeholders (enrolled parents, authorizers, etc.), these matters are beyond the purview of the current study. Rather, I pursue a topic modeling strategy with the goal of separating out "signals" of educational ideologies from other aspects of the charter school organizational form.

Topic models may be either single-membership or mixed-membership in structure: the former assumes the content of each document is drawn from a single topic, and thus a single distribution of words; the latter models documents as mixtures of topics, representing each document as a matrix of topic proportions, with its words drawn from these topics. In mixed-membership models—the focus of this paper—each word may appear in multiple topics, and a given word's meaning is determined (by the analyst) relative to every other word in the topic. Words and topics are assumed to be "exchangeable", that is, random variables derived from fixed conditional distributions (De Finetti 1975). This approach assumes that words in each document are generated by first choosing a topic and then choosing a word from that topic. Thus, topic models view words in documents as selected from a distribution of topics; each document contains multiple topics, and topics mediate between words and documents.

Mixed-membership topic models represent documents as a matrix of probabilities of words in documents and infers a set of topics that splits the word-document relationship. They discriminate themes based on word co-occurrences: words that occur together consistently across texts are more likely to be captured within a shared topic, the assumption being that these words convey similar shades of meaning. More particularly, such topic models produce lists of weighted words, which represent topics: for topics, higher weighted words more strongly reflect their shared meaning; for documents, higher weighted topics reflect interpretable thematic patterns. That is, such topic models recognize polysemy in language, for words may convey multiple meanings as captured in multiple topics. This linguistic structure aligns with how sociologists tend to think about themes in language, an intuitive advantage of topic modeling in addition to the reliable interpretability of the themes generated (DiMaggio et al. 2013).

The most significant user-selected parameter for topic models is the number of topics (k). With all such preprocessing and parametric options, each permutation is likely to produce a different grouping of the corpus (Grimmer and King 2011), for the resulting posterior distributions have multiple local modes amenable to heterogeneous solutions (e.g., Chuang et al. 2015; Sontag and Roy 2009). However, there is no correct number of topics and no fixed, objective diagnostic to derive this parameter; rather, the decision is highly application-specific and largely hinges on the degree of granularity desired by the researcher (Roberts et al. 2014). Thus, consensus among practitioners is that model output should be judged by substantive utility and interpretability rather than any "objective" measures (Blei 2012; DiMaggio 2015).

In mixed-membership topic models, document generation is modeled as three steps (Roberts et al. 2014:1066–67). First, the topic distribution (θ_d) is drawn from a global prior, such as a Poisson distribution (Blei et al. 2003).[8] The simplest and most popular topic modeling algorithm, Latent Dirichlet Allocation (LDA), assumes that the topic proportions (θ_d) are drawn from a Dirichlet distribution.[9] Then, for each (yet unassigned) word n in a document, a topic for that word ($z_{d,n}$) is drawn from a multinomial distribution based on that document's topic distribution. Finally, the observed word ($w_{d,n}$) is drawn from a distribution over the vocabulary for the chosen topic (β_k). Thus, the probability of each word in a document is modeled as the product of word probabilities within a given topic, $\varphi^{(k)} = P(w_n \mid z_n = k)$, and topic probabilities within a given document, $\theta^{(d)} = P(z_n = k \mid D = d)$. This gives (Fligstein, Stuart Brundage, and Schultz 2017:888):

$$P(w_n \mid D = d) = \sum_{k=1}^{K} P(w_n \mid z_n = k)\, P(z_n = k \mid D = d)$$

Structural Topic Models. Innovating on the topic model architecture just described, I use R's `stm` package to implement STM, a recently developed topic modeling algorithm that incorporates document-level information or metadata (here, a given school's socio-demographics) to more effectively identify latent themes. An increasingly popular method of automated text analysis in social sciences, STM has been applied to varied contexts, recently including feminist social movements (Nelson 2017), environmental sociology (Bohr and Dunlap 2018), and advocacy organizations (Bail, Brown, and Mann 2017). It also boasts a growing library of community-supported packages to assist interpretation and visualization.[10]

For technical notes on STM, see Appendix B. For an illustration of STM, see Figure 5.1.[11]

[FIGURE 5.1 ABOUT HERE]

To prepare the corpus for STM, I used the `prepDocuments` function built into R's `stm` package to lower case all words and remove punctuation and numbers. In addition, as is common

[8] Following convention in the topic modeling literature, d represents documents, w represents words, z represents topic assignments for each word, and k represents topics.
[9] Also known as a multivariate beta distribution, the Dirichlet distribution is the conjugate prior of the multinomial distribution. This makes it a computationally convenient distribution for topic proportions, because the word-topic distribution ($\varphi^{(k)}$) and topic-document distribution ($\theta^{(d)}$) can be updated without altering their multinomial form. Moreover, the Dirichlet distribution is commonly used in hierarchical clustering models of language (Blei, Ng, and Jordan 2003).
[10] See https://www.structuraltopicmodel.com/ for an up-to-date list of such packages as well as methods papers and published applications of STM.
[11] For a technical diagram of STM, see Roberts et al. 2013:2.

when preprocessing text for topic models, I stemmed all words with the commonly used Porter stemming algorithm (Porter 1980) to maximize semantic information by collapsing word forms (e.g., "runs" and "running" both reduce to "run"). I also removed stopwords using the SMART stopword list (Salton 1971) and removed rare words—that is, those that occur fewer than 30 times across the corpus.[12] Removing rare words resulting in dropping 130,344 out of 138,104 terms and 335,926 out of 2,971,949 words—leaving a total of 4,953 documents, 7,760 terms, and 2,636,022 words in the charter school websites corpus.

I included a range of topical prevalence covariates (see Appendix B) when implementing STM. The four key variables were socio-demographic proportions of: students of color in the schools, students receiving free or reduced-price lunch (a proxy for school poverty), people of color in the school district, and people in poverty in the school district. I also included grade range (dummies indicating whether a primary, middle, or high school), age of school (in 2015-16 school year; logged), number of students (logged), an indicator for urban locale status (1 if a school is located in a principal city and urbanized area, 0 otherwise), and the % of a website pages that are PDFs (e.g., student handbooks, charter applications)—which tend to be longer and more procedural in tone than other web pages and thus influence the language used.

To select the number of topics, I assess interpretability and granularity through iterative evaluation of alternative models as well by convenient metrics and functions the `stm` package provides to guide the user in selecting *k*. Such measures include likelihood of words held-out from documents; residuals of actual compared to predicted values; and lower bound of marginal likelihood. Perhaps the most useful metric is semantic coherence, which assesses consistent co-occurrence of topic words across topics and is known to track with expert evaluations (Mimno et al. 2011)—unlike comparable, predictive measures such as model perplexity, which is *negatively* correlated with human readings (Chang et al. 2009).[13]

By using the `searchk` tool in `stm` to compare possible values of *k* from 5 to 100, I find with my corpus that held-out likelihood and lower bound increase monotonically with *k*, while residuals decrease monotonically with *k*. While this pattern may naively point to a high value for *k*, the opposite indication is delivered by semantic coherence. Coherence sinks sharply until about 80 topics and then more gradually, encountering several local peaks or "elbows" from 19 to 31 topics before decreasing steadily. Indeed, outside of $k<10$, coherence attains its highest values in the 19 to 31 range. Accordingly, I ran four topic models with 19, 24, 28, and 31 topics and compared their results.

I find that with 19 and 24 topics, key themes connoting educational ideology become collapsed. For instance, the topics (see Table 5.1) I below label "Inquiry-Based Learning" (topic 7) and "Classical/Liberal Arts" (topic 26)—which are distinct with 28 or 31 topics—merge into one with 19 or 24 topics, with words like "classic", "scienc", and "art" distinguishing the

---

[12] I arrived at this relatively high threshold after finding more tolerant minimums (e.g., 5 or even 20 times) inadequate to the task of removing irrelevant language (e.g., rare names like "Johannes", typos like "schoool") that made inferring topics' meanings more difficult.

[13] However, coherence inflates artificially when *k* is low, due to the co-occurrence of frequent words—necessitating a complementary measure to offset this weakness. Filling this role is word-topic exclusivity, or the probability that a given word belongs to non-overlapping topics (Bischof and Airoldi 2012; Eisenstein, Ahmed, and Xing 2011): exclusivity is low in cases where overly similar topics heighten coherence (such as with few topics).

resulting hybrid topic. In contrast, with 31 topics I find that similar themes begin to fragment into overly specific word lists. For instance, in one topic “scienc”, “art”, “lab”, and “kindergarten” suggests environments designed for active learning; in another, the words “credit”, “transcript”, “biolog”, and “english” connote course credit. Balanced between collapsing themes and overly sharpening them is $k = 28$—my choice for the models in this study. However, I emphasize that all four options for k produce mostly similar clusters of topics and several consistent topics, as well—especially the themes on course requirements and online access, interface, and administration. Thus, while I choose 28 topics for purposes of semantic coherence and granularity, I anticipate qualitatively similar results across this space of parameter choices.

Dictionary methods. By counting the frequency of terms in the IBL *dictionary*—that is, a list of terms in an overarching category (Stone, Dunphy, and Smith 1966)—I measure the emphasis on IBL within each school’s identity. In dictionary methods, researchers develop a list of words connected to a category or concept of interest, then count instances of these words in a sample of texts for purposes of categorization (Grimmer and Stewart 2013).

Dictionary approaches rely on the classic linguistic assumption that language reflects culture: frequent words reflect the cognitive categories most on the author’s mind, while rare words are cognitively peripheral or alien (Sapir 1949; Whorf 1940). Thus, analyzing word frequencies on websites reveal how central is an ideology to a charter school’s identity: to the extent that a website uses the concepts in an ideology’s dictionary, the school identifies with that ideology.

Each dictionary is restricted in application to the ecological context in which it was generated and validated (Grimmer and Stewart 2013; Nelson et al. 2018). Indeed, because both researchers and text data are socially embedded, the significance of any set of terms shifts across sociolinguistic (Louwerse 2004) and sociocultural contexts (Henrich, Heine, and Norenzayan 2010). For instance, applying the Harvard General Inquirer (Stone et al. 1966) to classify negative tone in corporate earnings reports leads to serious misclassifications (Loughran and McDonald 2011). A classic example of applying a dictionary beyond its target domain, consequences include both many false positives (e.g., the term vice characterizes an executive rather than immorality, while tire is not negative in the context of automobile industry reports) and false negatives (e.g., litigation and unanticipated don’t appear in the Harvard list; Loughran and McDonald 2011). Moreover, language use changes with communication technology: for instance, social media incubates many context-specific expressions of tone (e.g., emoticons, hashtags), and as a result formal linguistic features (e.g., parts of speech) poorly predict sentiment (Kouloumpis, Wilson, and Moore 2011).

Despite the importance of domain-specific dictionary development and validation, many scholars continue to subjectively derive dictionaries from content knowledge or inspection of data (Egami et al. 2018) or rely on (supposedly) universal sentiment dictionaries developed and validated using large, non-specific text corpora (especially LIWC: Pennebaker, Francis, and Booth 2001; but also PANAS-X: Watson and Clark 1999). Scholars do so because dictionary development and validation is time- and resource-intensive--difficulties compounded by the absence of universal dictionary validation procedures (but see Grimmer and Stewart 2013) and the inability of manual content analysis to reproduce granular dictionary-derived measures (Krosnick 1999). And yet foregoing dictionary refinement and validation can obscure causal relationships: it’s often unclear which words drive statistical relationships, and relying on fixed (or even black-box) categories can conceal unexpected associations (Schwartz and Ungar 2015:81).

As such, in social research it is also common to develop idiosyncratic dictionaries suited to a specific measurement task and analytical domain, such as economic interest (Enns et al. 2016) or political ideology (Graham, Haidt, and Nosek 2009). Such dictionaries may be more hand-driven (Schwartz and Ungar 2015), or developed by experts from domain knowledge, literature review, and/or identifying synonyms; or more data-driven, or developed by means like crowd-sourcing (e.g., Benoit et al. 2016) or identifying distinguishing words.

I operationalize the emphasis on the IBL ideology as the ratio of the number of times a concept from the dictionary appears on a given website divided by the total number of words on that website (to account for varying lengths of websites). Ideological emphasis thus has a range of [0,1], where 0 indicates that *none* of the website's words are concepts from the ideology and 1 indicates that *all* of the website's words are concepts from the ideology (see Appendix C for examples of charter school websites with high and low IBL emphasis). Given that term counts for both IBL and word totals are skewed right (see Table 4.1 and Figure 4.2) to a degree that could change with website length, I account for possible bias by taking the log of each measure. I thus calculate emphasis as follows:[14]

$$Emphasis = \frac{\log(\#\text{ inquiry terms})}{\log(\#\text{ total terms})}$$

To my knowledge, no dictionary has been developed for a context like school websites; accordingly, I create and validate an IBL dictionary specific to my web corpus. Specifically, I construct word embedding models from charter school websites, and then I use these to iteratively expand a set of seed terms into a longer dictionary of words and phrases that have similar meaning as the core concepts (see below for explanation of this method).

Word embeddings. Word embeddings map words onto a high-dimensional vector space and represent semantic relations between words as geometric relations in space (Mikolov, Sutskever, et al. 2013). Such techniques are frequently used in digital humanities research; for example, to trace the shifting meaning of "gay" over the twentieth century from a position near "dapper" and "cheerful" to one close to "lesbian" and "homosexual" (Kulkarni et al. 2015). As an inductive approach, word embeddings are similar to other inductive quantitative methods commonly used in the social sciences, such as factor analysis, latent factor analysis, or multidimensional scaling. Yet while similar inductive textual approaches limit the scale of analysis to ease interpretability—for instance, semantic network analysis simplifies associations in large corpora, while topic modeling underweights most word-topic and topic-document associations—word embeddings offer the advantage of efficiently modeling cultural dimensions based on semantic relations among *all* words. Word embeddings are becoming more common in the social sciences; for example, to analyze associations between basic cultural categories such as gender, race, and status/wealth (Kozlowski, Taddy, and Evans 2019).

A word's position in vector space is based on the context that it shares with other words in the focal text. Words that share many contexts (i.e., words that are frequently collocated with the

[14] To avoid dropping from my models schools with no IBL terms on their websites—otherwise producing an undefined numerator from `log(0)`—I add `1` to both the # IBL terms and the # total terms when implementing this formula.

same other words) are positioned near each other in vector space, and words that have very different contexts (i.e., words that are collocated with different other words) are positioned far apart. In other words, words that are positioned near each other in vector space share similar meanings, so vector space can be understood as semantic space. Importantly, this relational mapping captures commonalities in words' local contexts, rather than collocation alone. This large-scale mapping of contexts encodes word embeddings with underlying cultural meanings, rather than strictly on-the-ground observable patterns.

Mechanically, word embeddings assess word associations using 'word context windows', indicating the number of words (typically 5-12) on either side of focal word w to consider as connected to w. Formally, for a series of training words $w_1, w_2, w_3, \ldots, w_T$, the goal of word embeddings to maximize the average log probability of predicting w_{t+j} given w_t (Mikolov, Sutskever, et al. 2013:2):

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p\left(w_{t+j}\middle|w_t\right) \tag{1}$$

where w indicates a word in the sentence, t is the iterator over T training words in the sample sentence, c is the word context window size, and j is a number between $-c$ and c (excluding zero) that indicates the distance in words from focal word w_t to a word within its context, w_{t+j}.

The high number of dimensions in the vector space (typically 100-300) means that Euclidean or straight-line distances between vectors cannot be calculated (Kozlowski et al. 2019:9). Instead, the distances between word vectors are established via cosine similarity, which measures the angle between vectors such that a score of 0 indicates perfect independence (orthogonality or 90 degrees between vectors) and 1 indicates perfect similarity (parallelity or 0 degrees between vectors). See Appendix B for additional technical notes on word embeddings.

Development of IBL dictionary. To borrow from the rich literature on IBL in educational psychology, I began with five seed terms taken directly from the subtitle of a seminal article: "The failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching" (Kirschner, Sweller, and Clark 2006). Thus, my five seed terms are: "inquiry-based", "problem-based", "discovery-based", "experiential", and "constructivist".[15] Indeed, though conceptually distinct (e.g., Steffe and Gale 1995), empirical research generally treats these terms as synonyms. My word embeddings support this claim: the average cosine similarity of these five terms is high at 0.70.

In a blended data- and hand-driven procedure (Nelson 2017; Schwartz and Ungar 2015), I used the `most_similar()` method in Python's `word2vec` module (Řehůřek and Sojka 2010) to identify word vectors positioned in the vector space near the seed terms, indicating semantic similarity. I then applied two manual filters to focus the dictionary. First, I removed terms conceptually distinct (in my eyes) from IBL. For example, I removed the terms "strengths-based", "small-group", "explorations", and "self-motivated" because these teaching methods and desirable traits can be taught without IBL—regardless of their cosine similarities to the IBL seed terms (0.76, 0.59, 0.56, and 0.51, respectively). Second, I removed terms that occurred rarely in

[15] I changed the word "discovery" to "discovery-based" because unlike the celebratory former, in my web corpus the latter is reliably pedagogical in meaning.

the corpus, such as “sustained inquiry” (7 counts) and “multimodal” (14 counts), but kept any obvious synonyms to the seed terms, such as “problem-centered” (12 counts, but identical to “problem-based”). These filters ensured that new terms were theoretically relevant; for this reason, manual checks are a common means of validating dictionaries built with word embeddings (e.g., Sivak and Smirnov 2019; see “robustness checks” section for more on dictionary validation). For visualization of the full dictionary, see Figure 4.6; for per-term counts and cosine similarities (between a given term and the seed terms), see Table 4.2.

[FIGURE 4.6 ABOUT HERE]

[TABLE 4.2 ABOUT HERE]

Mixed-effects models. I next estimate mixed-effects linear models (which have both fixed and random effects) in STATA 15 for dependent variables at the school level. Given that state policy (e.g., Bodine et al. 2008; Finnigan 2007), school district practices (e.g., Paino 2018), and CMO structure (e.g., Furgeson et al. 2012) have significant, convergent impacts on charter school practices and outcomes, ignoring the nesting of charter schools (level-1) in states, school districts, and CMOs (level-2) by pooling all observations would violate the assumption of independence required for ordinary least squares (OLS) regression and thus bias the results.[16]

Furthermore, these are cross-classified data in that each CMO could have schools in multiple states, and each state could be home to multiple CMOs. These data are also hierarchically clustered (Goldstein 1987) in that school districts are nested within states. To accommodate this data structure, my mixed-effects models could estimate crossed random effects for state and CMO as well as random effects for school districts nested in states. Such models take the following form:

$$y_{isdc} = \beta_0 + \beta' x_{isdc} + \zeta_s + \zeta_d + \zeta_c + \varepsilon_{isdc} \quad (2)$$

where *i* represents schools, *s* represents states, *d* represents school districts, *c* represents CMOs, y_{isdc} is the emphasis on IBL, β_0 is the intercept, x_{isdc} is a vector of explanatory and control variables, ζ_s is the random effect for state *s*, ζ_d is the random effect for school district *d,* ζ_c is the random effect for CMO *c*, and ε_{isdc} is the error term. The first random effect captures unobserved factors that might shape each state’s effect on school ideology; the second captures those factors that might affect ideology within each school district; and the third captures those that might affect ideology in each CMO. Thus, in this full model, controls at the levels of state, school district, and CMO are not necessary.

However, because estimating all three random effects (CMO, state, school district) in all three models (predicting IBL, school poverty, or school race) is inefficient and ignores the extent of nesting, I assessed nestedness by measuring the Intraclass Correlation Coefficient (ICC) or

[16] While modeling with fixed effects (FE) controls for level 2 influences (e.g., state policy, CMO size) and thus eliminates unobserved sources of heterogeneity (Schneider et al. 2007), pure FE models give poor estimates for low-sample groups and control away all across-state, across-district, and across-CMO variation in outcomes. Mixed-effects models resolve these concerns through partial pooling—estimates for states and CMOs with fewer observations are based partially on the estimates for larger such groups.

rho (ρ) for each level through preliminary models. By measuring the proportion of residual variance in the outcome explained by the non-independence (nesting) of units in each level, the ICC indicates whether a random effect for that level is necessary. I thus discovered that IBL is nested in CMO ($\rho = 0.31$) but not state ($\rho = 0.04$) or school district ($\rho = 0.09$); that school poverty is nested in school district ($\rho = 0.23$) much more than state ($\rho = 0.12$) or CMO ($\rho = 0.14$); and that school ethnicity is nested in school district ($\rho = 0.33$) and state ($\rho = 0.26$), but not significantly in CMO ($\rho = 0.12$). Accordingly, I nested my IBL models by CMO, my school poverty models by school district, and my school ethnicity models by state and school district. I report revised, model-specific ICC values below.

I include school-level controls for grade range (dummies indicating whether a primary, middle, or high school), age of school (in 2015-16 school year; logged), number of students (logged), and an indicator for urban locale status (1 if a school is located in a principal city and urbanized area, 0 otherwise). In models with IBL, I also include a control for the % of a website pages that are PDFs (e.g., student handbooks, charter applications), which tend to be longer and more procedural in tone than other web pages and thus influence the language used. Lastly, in models with academic proficiency rates, I include measures of data "blurring" by the U.S. Department of Education to protect the identities of (especially small) student groups. These measures scale with data precision: possible values are 1 (percentiles reported, e.g. 95% academic proficiency for a school; the most precise), 5 (quintiles reported, e.g. 90-95%), 10 (deciles reported, e.g. 80-90%), 20 (ventiles reported, e.g. 60-80%), and 50 (medians reported, e.g. 50-100%; the least precise).

Modeling approach

In order to account for alternative influences on IBL and charter school enrollments, I build two sets of models: one predicting IBL emphasis, and the other predicting proportions of low-income students and (separately) students of color.

Regarding the first set of models, the social context within which charter schools seek to secure resources is not confined to those already enrolled in the school. They must appeal also to potential 'clients' within the school district—the political and administrative arena in which policies are enacted (e.g., Finnigan 2007), parents and others exert influence (e.g., Preston et al. 2012), and schools compete for students and favor (e.g., Arum 1996). Indeed, just as sociodemographic factors drive parents to sort themselves into schools, so do they drive schools to sort themselves into districts—an influence all the more pronounced given persistent residential segregation by race and income (e.g., Massey, Rothwell, and Domina 2009; Reardon, Townsend, and Fox 2017). Thus, social context in the school district—rather than the school—represents an alternative mechanism for my hypothesis. Accordingly, I analyze the relationships between IBL and (as independent variables) both school and school district socioeconomic and ethnic composition.

Regarding the second set of models, school choice scholars often assume that parents respond rationally to their educational options, choosing the highest quality schools—that is, those that perform best on standardized tests—available for their children (e.g., Epple, Figlio, and Romano 2004; Hanushek et al. 2007). Moreover, academic quality is generally reported at the top of parents' educational values—with school safety, extracurriculars, and moral instruction taking precedence only in select circumstances (for reviews, see Erickson 2017; Posey-Maddox, Kimelberg, and Cucchiara 2014). But I argue that parents' school choices are not driven solely by objective signals of academic quality; rather, educational ideology plays a role in shaping enrollment patterns. To support this claim, I analyze the relationships between school socioeconomic and ethnic composition and (as independent variables) IBL and academic quality.

Recent empirical research supports this challenge to the primacy of academic quality. Scholars may have exaggerated the influence of objective academic quality on parents' school preferences, which are better explained by peer quality (the performance of the existing student body; Abdulkadiroglu et al. 2019; Rothstein 2006). Indeed, much research shows that "school quality" is not an objective signal, but rather is socially constructed in ways that reflect hierarchies of race and class. Even high-status parents—who may possess better information on school composition, achievement, etc. than do low-status parents (e.g., Teske, Fitzpatrick, and Kaplan 2006; Yettick 2016)—assess school quality through their social networks, rather than relying on objective test score data or first-hand observation (Holme 2002). And even those white, middle-class parents committed to reducing inequalities and valuing diverse schools send their children to the whitest, best-funded public schools they can (e.g., Reay et al. 2008; Roda and Wells 2013).

The first set of mixed linear models feature a lagged dependent variable: IBL was captured via web-crawling in June 2018, while the sociodemographic predictors were measured in 2015-16 (at school level) or 2012-16 (at school district level). This time lapse between measures strengthens the argument that sociodemographics influences educational ideology, because this means potential founders have over two years to establish new schools or retune their educational approach to respond to local demands. In predicting school sociodemographics, the second set of models similarly uses lagged academic proficiency scores (measured in 2014-15), allowing parents and resource providers time to observe and react to objective signals of school quality. However, the second set also includes educational ideology, a covariate measured several years prior to the outcome (June 2018 compared to 2015-16)—making the second set of models a more conservative test of my hypothesis.

CHAPTER 5: RESULTS OF STRUCTURAL TOPIC MODELING

Distinctive words by race

Before considering the results of Structural Topic Models, for purposes of comparison figure 5.2 shows the results of a simpler method: detecting via log-odds ratios the most distinctive words for schools below or above the median proportion of white students.

[FIGURE 5.2 ABOUT HERE]

Notable words more commonly used in white-serving schools include “homeschool”, “montessori” (a holistic educational approach), and “classic” (as in a “Classical” or well-rounded curriculum), each signaling the upper-class parenting style of concerted cultivation discussed above—in particular, cultivation of individual talents and child-centered pedagogy. Similarly, terms such as “band”, “pe”, and “virtual” indicate access to non-core subjects, extracurriculars, and adaptive curricula, while “fee”, “cell”, and “fundraising” reflect parental involvement or resources. In contrast, words more distinctive of schools enrolling students of color include “harmony”, “justice”, “neighborhood”, and “transform”, indicating an ethos of personal responsibility for lifting communities and overcoming disadvantage. Likewise, the words “scholar”, “collegiate”, and “prep” clearly indicate a focus on academic achievement and the more traditional learning focus of the working-class model of accomplishing natural growth (see above). Finally, “ace”, “aspire”, and “idea” are all names of sizable CMOs largely serving urban and minority communities, reflecting growing corporate influence in the charter school sector.

While comparing distinctive words along this single dimension (proportion white students) is instructive, it provides a limited lens into the niche-specific language used by schools situated within and seeking to appeal to socio-demographic niches. To capture such linguistic intersectionality precisely—that is, with estimates of statistical confidence—I use Structural Topic Modeling.

Overview of topics

Figure 5.3 shows the general results of the 28-topic Structural Topic Model: the frequency and top four words per topic as ranked by the simplified frequency-exclusivity (FREX) metric, which balances the probability of a word occurring in a given topic with the exclusivity of the word to that topic (Roberts et al. 2014, 2019). This makes for easier interpretation than when using raw word probabilities (i.e., the probability of finding a word conditional on the topic chosen). Specifically, FREX is derived from the harmonic mean of a given word’s frequency and exclusivity, such that neither rare words (which by default tend to have high exclusivity) nor frequent words (which typically contribute little to a topic’s meaning) have high scores (Airoldi and Bischof 2016; Bischof and Airoldi 2012). This logic is similar to that described above for using both coherence with exclusivity to evaluate topics. Moreover, these are similar advantages to the *term frequency by inverse document frequency* weighting scheme popular in information retrieval (Salton and Buckley 1988).

[FIGURE 5.3 ABOUT HERE]

A variety of topics (each labeled in the next figure) are evident, with distinctive words ranging from the more academic (e.g., “montessori” in topic 13 and “project” in topic 7), to the more digital (e.g., “site” in topic 10 and “edlio”[17] in topic 9), to the more organizational in focus (e.g., “week” in topic 14 and “suspens” in topic 27). I label each topic inductively based on my review of the most probable and distinctive word stems as well as representative documents. I now describe each topic and explain its label (all topics are labelled in the next figure); I also overview patterns in topics’ relationships.

The most common topic is by far topic 18, which I label the Standards and Assessment topic based on its distinctive word stems (e.g., “assess”, “achiev”).

The second most common is topic 10, which has several web-related distinctive terms (e.g., “map”, “site”, “link”) but emphasizes methods of contact and social media (top words include “comment”, “twitter”, and “fax”). Thus, I label this the Communication topic. However, other topics share this emphasis on online experience and contact—including the next most common, topic 5, with a number of similarly digital-focused distinctive terms (e.g., “click”, “download”, “portal”), but many others more administrative in subject matter (e.g., “suppli”, “tour”, “registr”). As such, I label topic 5 as Web Administration. Similar but less frequent topics includes topic 8, labelled Web Interface (distinctive words include “web”, “enabl”, “select”); topic 9, or Online Access (e.g., “verif”, “authent”, and “sorri”—meaning an apology for access issues or locked content); and topic 24, or Terms of Agreement (e.g., “user”, “agre”, “privaci”).

The fourth most common topic is 13, which I label Holistic Education given the emphasis on progressive educational models (e.g., “montessori”, “waldorf”), methods (e.g., “garden”, “emot”—meaning socio-emotional learning), and values (e.g., “beauti”, “children”—suggesting innocence rather than rigor). The next most common is topic 4, which also emphasizes academic content or educational process: labeled Conversation, it includes informal or narrative language (e.g., “thing”, “hope”, “lot”) and suggests storytelling, social self-expression, or students’ perspectives. The next distinctly academic topics are 1, which I label as Course Requirements (e.g., “cours”, “credit”, “transcript”); 7, or Inquiry-Based Learning (e.g., “art”, “project”, “music”); as well as the much less common topic 26, or Classical/Liberal Arts education (e.g., “classic”, “grammar”, “latin”).

Also toward the top of Figure 5.3 is topic 12 or Faculty and Staff, composed of language indicating staff qualifications (e.g., “bachelor”, “degre”) and biographies (e.g., “born”, “bio”, “passion”). The next most frequent topic also relates to school organization and its functions: topic 14 or Schedules/Calendars, which describes time (e.g., “week”, “friday”, “jun”), routine activities (e.g., “book”, “remind”, “dismiss”), and school levels (e.g., “grade”, “fourth”). Less frequent topics of similarly organizational subject matter include 15, or Dress Code, which regulates clothing and appearance (e.g., “worn”, “uniform”, “permit”); 16, or School Governance, describing the school board and related official activities (e.g., “agenda”, “minut”, “board”); 22, or Meals (e.g., “chicken”, “sandwich”, “oven”); and 27, or Discipline, which covers behavioral standards and disciplinary protocols (e.g., “shall”, “suspens”, “harass”, “appropri”).

---

[17] Edlio is a school website design service commonly used by charter and traditional public schools; see https://edlio.com/.

To compliment this (non-exhaustive) overview of the topics, I next test the associations between each topic and key socio-demographic variables: school and district poverty and ethnic composition.

Topics by race and class

To test the associations between topics and school and district socio-demographics, I use the `EstimateEffect` function within the `stm` package in R, which uses linear regression to predict each document's proportion about a topic conditional on document characteristics and estimates uncertainty based on the average covariance matrix of the STM model itself (Roberts et al. 2019). I compute these regressions with and without school-level controls; the latter results are close to raw correlations, while the former are more robust to alternative explanations. Controls overlap with the topical prevalence covariates described above, except for the key socio-demographic variables; thus, I control for grade level, age, size, urban status, and % PDF pages. Table 5.1 shows the results of these regressions.

[TABLE 5.1 ABOUT HERE]

Many topics have significant relationships with race and class—and many are robust to controls. Another notable trend is the consistency of the above associations: those topics with significant negative associations with a given covariate tend to also have negative associations with other covariates, regardless of the level (school or district) and whether controls are included. For instance, topics 1 and 4 have consistent, strong negative associations with school or district poverty and race. The converse is also true for positive associations—such as the consistently strong, positive associations between each of topics 11 or 17 and school or district poverty and race.

For visualization of selected topic proportions with sociodemographics, see Figures 5.5-5.9 (below).

Close inspection of these topic/covariate associations—as well as their thematic groupings and distinctive words—suggests distinct clusters of topics in the STM model, which I now discuss.

Clusters of topics

I identify six clusters of topics, each of which shares an overarching theme and prevailing associations with race and poverty. The largest and most conspicuous such cluster is *Academics*, including the topics I label as Course Requirements (topic 1), Conversation (4), Inquiry-Based Learning (7), Holistic Education (13), and Classical/Liberal Arts (26). Each of these topics has a strong ($p<0.001$) negative association with people of color and poverty at the school and district levels (i.e., a positive association with white and affluent schools and districts), and these relationships are largely robust to controls. Contrasting the *Academics* cluster are two others strongly associated instead with students of color and poverty—but mixed (though mostly expected) relationships with district demographics, and more sensitivity to controls. These are the *Standards/ College-Bound* cluster, with topics on College Prep (17), Standards/Assessment (18), and Diploma/Graduation (20); and the *Serving the Disadvantaged* cluster, containing topics on Athletics (2), Urban Locales (11), Virtual Education (19), and Services (23).

The clusters I call *Academics, Standards/College-Bound,* and *Serving the Disadvantaged* not only have strong, consistent associations within the cluster, but also they are each relevant to the central construct I seek to capture: educational ideology as a marker of niche-specific organizational identity. In other words, these clusters contain opposing, culturally situated claims as to what constitutes an ideal learning environment. For the *Academics* cluster (distinctive to white schools), education entails a coherent, humanities-rich course program (topics 1 and 26) that instills comfortable self-expression (topic 4) and appreciation of nature (topic 13) through creative, original projects (topic 7). For the *Standards/College-Bound* cluster (distinctive to schools serving students of color), education means a college-preparatory (topic 17), standards-aligned curriculum (topic 18) that ensures access to diplomas (20). Likewise, for the *Serving the Disadvantaged* cluster, schools should address the difficulties that poor or marginalized students face in traditional environments: by providing alternative pathways to college (e.g., athletics; topic 2), virtual learning (topic 19), and a range of services (e.g., nutrition, medical; topic 23), especially in urban-situated schools (topic 11). Each of these cluster's race/class associations and distinctive words per topic are shown in table 5.2, and said associations along with confidence intervals are visualized in figures 5.5-5.9.

[TABLE 5.2 ABOUT HERE]

[FIGURES 5.5-5.9 ABOUT HERE]

I also identify three other, less ideologically relevant clusters including *School Organization/ Management*, which contains a set of apparently heterogeneous topics: Faculty and Staff (12), Schedules/Calendars (14), Dress Code (15), School Governance (16), Meals (22), and Student Discipline (27). In addition to their common (though multifaceted) focus on running a school, however, these topics are also all associated with more white and less poor schools and districts—though most of these relationships (other than for topic 14) are inconsistent and generally not robust to controls. The remaining two clusters are *International*, containing the topics of Spanish Language (25) and International (28), which are associated with poor students and students of color (especially Hispanic students); and *Website*, with the topics of Web Administration (5), Web Interface (8), Online Access (9), Communication (10), and Terms of Agreement (24), which are weakly associated with white, affluent schools and districts (except topic 9, which goes in the opposite direction). Topics outside my cluster scheme are School Names and Places (3), with a scattershot of place names (e.g., "haven", "philadelphia", "hill") and no significant associations; Civil Rights (6), which altogether lacks significant associations; and Parent Involvement (21; distinctive words include "volunt", "join", and "trip"), which doesn't fit the focus on programs and services in the *Serving the Disadvantaged* cluster but has similar associations with poverty and people of color.

Discussion

These results are a nuanced first look at the diversity of themes in my corpus. As expected of website text data, several topics in what I call the *Website* cluster convey information specific to the online user experience: where to click, how to access newsletters, the school's social media handles, etc. Still other topics communicate requisite information on *School Organization*: weekly schedules, dress codes, faculty and staff qualifications, etc. Others contain information like the school's name, policies towards discrimination, opportunities for parental involvement, or terms in foreign languages or on foreign lands.

Most important, however, are those topics that convey race- and class-specific educational approaches. The *Academics*, *Standards/College-Bound*, and *Serving the Disadvantaged* clusters of topics share the strongest, most consistent associations with race and class—reinforcing their thematic coherence and relevance to educational ideology. Of these, the *Academics* cluster is particularly distinguished for its significant race- and class-specificity and practical implications for teaching through conversation, time in nature, project-based learning, and humanities courses.

Beyond this unsupervised text analysis, the next steps are finer-grained examination of these discovered themes and additional computational methods for pattern confirmation (Nelson 2017). Indeed, such measures are necessary to address two key limitations in this analysis. First, STM does not incorporate local semantic (i.e., sentence-level) information, e.g. the co-occurrence of certain words within "windows" of words. This is due to the "bag of words" assumption inherent to topic modeling—regardless of STM's incorporation of document-level information when generating distributions. Second, the method of linear regression used to estimate relationships between topics and school socio-demographics lacks the complexity to account for the nesting of schools in school districts, states, and CMOs. To account for these unobserved correlations among units, a more sophisticated modeling approach is necessary.

The next chapter addresses both these limitations: first, by using sentence-level semantic information (through word embedding models) to expand a list of seed terms; and second, by implementing mixed-effects linear models to account for nesting of schools in multiple, crossed levels. To provide a list of seed terms, I draw on a topic from the *Academics* cluster: Inquiry-Based Learning.

CHAPTER 6: RESULTS OF DICTIONARIES AND MIXED LINEAR MODELS

Correlations

The correlations between variables (see Table 4.1) and their visualizations (see Figures 6.1 – 6.8) support my analytic approach. IBL correlates with demographic contexts in the directions predicted: IBL is negatively correlated with school percentage poverty and percentage students of color and school district percentage poverty and percentage people of color (-0.18, -0.17, -0.11, and -0.06, respectively), and all of these correlations are significant at the $p<0.05$ level. Moreover, academic proficiency is negatively correlated with poverty and percent students of color at both the school and district levels, suggesting that academic quality may be an alternative explanation for charter school enrollment patterns. These correlations tentatively support my hypothesis—though they are subject to confounding and do not disentangle causal relationships.

[FIGURES 6.1 – 6.8 ABOUT HERE]

In addition, the correlations between the demographic variables are positive and strong, ranging from 0.27 (between school district percentage people of color and school percentage poverty) and 0.62 (between school district percentage people of color and school percentage students of color). Due to this high correlation among sociodemographic variables, the direct relationship of each with the outcome would be muddled if all were included in a single model; thus, I estimate the effect of each of these independent variables separately.

Mixed-effects models

The mixed-effects linear models regressing IBL emphasis on school and school district poverty and race—plus school controls—are shown in Table 6.1. These findings support my hypothesis, for school and district percentage poverty and percentage nonwhite all have statistically significant, negative relationships with IBL emphasis in their respective models (see Figure 5.9). Moreover, the effect sizes are considerable: each standard deviation increase in school poverty, school percentage students of color, district poverty, or district percentage people of color is associated in their respective models with a change in IBL emphasis of -1.48, -1.63, -1.15, and -0.418 standard deviations, respectively. Thus, not only is the effect of school poverty and ethnic composition on IBL greater than for school district poverty and race, but also their models (1b and 1c) have a higher log-likelihood and lower AIC and BIC than the others (1d and 1e). This difference is most dramatic for school district race, which unlike school district poverty has a small beta coefficient as well (-0.03 vs. -0.21). As such, these findings suggest that sociodemographics—and especially poverty—at the level of school, rather than the school district, have the stronger, more negative relationship with IBL emphasis.

Furthermore, the ICC indicates that CMO membership explains an impressive 34% or so of the variation in IBL emphasis, and the conservative χ^2 test confirms the need in the model for random effects. Together, these findings suggest that schools sharing a CMO are significantly more alike in their IBL emphasis than are schools not sharing a CMO.

[FIGURE 6.9 ABOUT HERE]

[TABLE 6.1 ABOUT HERE]

The mixed-effects linear models regressing school poverty and school ethnicity on IBL emphasis and academic quality—plus school controls—are shown in Table 6.2. These findings support my hypothesis, for IBL emphasis has statistically significant, positive relationships with school poverty and ethnicity in their respective models. Furthermore, this effect does not disappear when academic proficiency is considered (see Figure 6.10)—though it does decrease in size, the effects of proficiency in reading/language arts and math also dip slightly. In the full models, each standard deviation increase in IBL emphasis is associated with a change in school poverty and percent students of color of -2.32 and -3.18, respectively. In contrast, the same unit change for reading/language arts proficiency is -4.60 and -4.45, and for math proficiency it is -0.741 and -0.901. Moreover, those models including academic proficiency have higher log likelihood and lower AIC and BIC measures than do the other models, underscoring the stronger impact of reading/language arts proficiency on enrollments. Nonetheless, the effect of IBL emphasis is greater than that of math proficiency, which is only marginally significant.

In these models, ICC measures indicate that school district membership explains about 37% of the variation in both school poverty and school ethnicity. And the conservative χ^2 test confirms the need for random effects, which are statistically significant ($p < 0.01$) at the CMO and school district levels. Together, these findings underscore the sociodemographic similarity of schools sharing a school district.

[FIGURE 6.10 ABOUT HERE]

[TABLE 6.2 ABOUT HERE]

To review, not only are demographic factors correlated with IBL emphasis in the predicted directions, but also these same directions hold with statistical significance in each model. School poverty and ethnicity and school district poverty all have strong relationships with IBL emphasis. Moreover, the significant relationship between IBL emphasis and school sociodemographics is robust to traditional measures of school quality. This effectively discounts the alternative explanation that charter school identity has no link with enrollment patterns once academic quality is considered. Finally, both ideology and sociodemographics are embedded in nested organizational contexts: significant variation in IBL is explained by CMO membership, while substantial variation in school class and race is explained by school district and state membership.

Robustness checks.[18] The impact of objective quality signals may be greater when using further lagged predictors, for then parents and resource providers have longer than one year to observe and react to these signals. To test this possibility, I replicated the above analyses using academic proficiency rates for 2013-14 instead of 2014-15. These results were not significantly

[18] For Stata logs of each of these robustness checks (including dictionary validation), see the replication repository at https://github.com/URAP-charter/sorting-schools-2020/. Scroll lower in the page for a guide that corresponds log files to the robustness checks described in the text.

different from those reported above. Similarly, because ideologically distinct schools' may attract students outside their district catchment area, the impact of school enrollment demographics may be better revealed when operationalized relative to their surrounding district. As such, I re-ran the analyses using not within-school demographic proportions (e.g., percent students of color) but instead differentials between the school and district demographics (e.g., percent students of color at school level subtracted from proportion people of color at district level). Math proficiency grew (from 0.059 to 0.108 in absolute value) and statistical significance (from $p<0.05$ to $p<0.001$) when predicting the school poverty differential, but the results were otherwise unchanged.

To test whether the above results were skewed by schools with less precise academic data (perhaps due to misspecification), I also re-ran these analyses using only those 1,982 schools with proficiency scores reported in percentiles (the most precise). While math proficiency changed sign and largely became statistically insignificant, the direction, approximate effect size, and ratio between IBL emphasis and academic proficiency measures were otherwise consistent with the above.

Outliers could have also skewed the above results. To test the possibility that the above results are biased by very large websites (100 web pages or longer) or small schools (10 students or less), I also replicated the above analyses with filtered data. Thus, I replicated once for data with very large websites removed and once for data with small schools removed. Although math proficiency became marginally statistically significant when large websites were removed, the results were otherwise unchanged.

Likewise, given that IBL is more common in schools with more affluent, white students, the above results may not represent schools in more poor, racially diverse areas. To evaluate whether my results are thus limited in geographic reach, I repeated the analyses after filtering to only those schools located in school districts with above-median disadvantage by three separate metrics: proportion in poverty, proportion people of color, or population density (a continuous measure of urban status). I analyzed each filtered data set separately, amounting to three additional batches of models. Except for declining significance of those variables used for filtering—e.g., school district poverty became insignificant for the data set filtered to only those schools in districts with above-median poverty—and marginal significance for math proficiency in some cases, all these results were consistent with the above.

Moreover, to test the possibility that inclusion of full nesting would change the results, I replicated all models using random effects for CMO, state, and school district. The results were essentially identical—except that math proficiency was more significant when predicting school poverty ($p < 0.000$ instead of $p < 0.05$).

Validation of IBL dictionary. Finally, the particular words chosen for the IBL dictionary may be driving the results. In particular, the term "hands-on" occurs 48,423 times—nearly seven times as frequently as the next most common word in the dictionary, "problem-solving", at 7,125 times. Accordingly, I test the robustness of the above findings by calculating Emphasis and running mixed linear models for three additional dictionaries: the five seed terms; a 20-term, theoretically narrow IBL dictionary including only synonyms of the seed terms (a subset of the full dictionary); and the full dictionary with the "hands-on" term removed, leaving 49 terms (see Table 4.2 for detail on these dictionaries).

The results of these additional models strongly support the results discussed above. When predicting IBL, schools are still nested in CMOs (though less so for the seed dictionary: $\rho =$

0.18) but not states or districts, and the only key covariate (i.e., school and school district race and poverty) that loses statistical significance compared to the above results is school district proportion people of color ($p < 0.226$) when using seed terms. Otherwise, the only change for key covariates when predicting IBL is a decrease in effect size proportional (though not linearly) to dictionary size. Specifically, beta coefficients are 3-4 times smaller when seed terms are used (except for school district race, which is even smaller but not significant); a quarter to a third smaller when the narrow, 20-term dictionary is used; and almost identical when only "hands-on" is removed. This shrinking of betas is to be expected given that total word count decreases proportional to dictionary length—though it continues to discriminate effectively between educational ideologies, as consistent statistical significance indicates.

When predicting school poverty and race, the only significant change is an increase in betas for IBL, while the betas for academic proficiency are static. Specifically, betas roughly double for the seed dictionary, increase 1.2-1.5 times for the narrow dictionary, and are (again) almost identical for the 49-term dictionary. The inverse relationship between dictionary size and IBL betas suggests that shorter, conceptually sharper dictionaries may capture the clearest signal of organizational identity, and thus better explain demographic outcomes.

CHAPTER 7: CONCLUSIONS

These results indicate that charter schools' identities are associated with class- and race-differentiated social contexts. Through inductive text-analytic methods, I discern in school websites clusters of ideological themes strongly associated with race and class at the school and school district levels. Schools serving white and affluent students and districts are more likely to emphasize course information, conversational language, holistic education, liberal arts, and IBL. In contrast, schools serving those in poverty or people of color place greater emphasis on college preparation, school accountability, and diploma attainment, as well as urban environments and services (e.g., health, nutrition, salaries) and special programs like athletics and virtual learning. With precise statistical models, I also find that charter schools present themselves to affluent and (especially) white communities in ways emphasizing IBL, and that schools emphasizing IBL have student enrollments that are more affluent and white—independent of objective measures of school quality.

These conclusions provide initial support for my theory that schools' self-presentation strategies—in particular, their educational ideologies—respond to race- and class-specific educational values and expectations (e.g., Erickson 2017; Zeehandelaar and Winkler 2013) and culturally distinct parenting styles (Lareau 2000, 2011). That is, charter schools present themselves differently—by virtue of their explicit educational ideologies—to different race and class niches (Carroll 1985; Lauen et al. 2015). And this relationship may be driven as much by parents selecting schools, evidenced by the relationship between school race and poverty and the educational ideology of IBL; as by schools selecting districts, evidenced by the relationship between district race and poverty and IBL.

Thus, I demonstrate strong associations between educational ideologies and socio-demographics, supporting previous research (Malkus and Hatfield 2017; McShane and Hatfield 2015) in showing that progressive education is especially popular among white schools, while traditionalist models prevail among schools serving students of color. Moreover, building on clustering models (King et al. 2011) and analyses of resource partitioning among charter school identities (Renzulli et al. 2015), I find that mission specialization intersects with target population. In particular, specialized academic programs and creativity- and arts-based education are favored by affluent or white schools, while specialized programs and resources are more common in schools serving poor students or students of color.

My research helps resolve the ambiguity pressing the public, academia, and policy circles over the relationship between charter schools, on one hand, and persistent racial and socioeconomic divisions in U.S. society, on the other. An imperative for research into educational stratification and school choice is to understand the extent to which charter schools' strategic identities—in appealing to parents' distinctive tastes—consolidate parents by race and class, amounting to a vector of segregation by race and class. Yet only first steps have been made. Some scholars document organizational pathways for charter school segregation (e.g., attracting students of color with culturally relevant programs: Fabricant and Fine 2012; selecting students based on performance: Adamson and Darling-Hammond 2016; Lacireno-Paquet et al. 2002)—but school choice research has not expanded this organizational account. Research finds considerable variation among charter school identities (e.g., King et al. 2011; Renzulli et al. 2015)—but stops short of analyzing their social contexts and impacts, leaving advocates room to claim that distinct organizational forms best meet the needs of distinct student groups (e.g., Malkus and Hatfield 2017; McShane and Hatfield 2015). Underlying this gap in the literature is a dearth of research at the nexus of organizations and education: organizational research is all but silent on recent educational transformations (Renzulli 2014), while educational sociology shows greater interest in school effects than in politics, curricula, or social contexts (Brint 2013).

Thus, no research to date has examined how organizational identities in charter schools—the most widespread school choice reform today (Berends 2015)—interact with the community social forces underlying parents' educational preferences (for a review, see Erickson 2017) to create charter school segregation. In beginning to fill this gap, my findings suggest (but cannot prove) that the ideological differentiation of charter schools reinforces social inequalities—rather than alleviates them, as educational reformers claim (e.g., Pendergrass and Kern 2017; Roth et al. 2017). The present study initiates a research program to examine evidence that organizationally differentiated identities attract socially differentiated 'clients', a mechanism by which parents may self-sort along dimensions of inequality such as race and class. Multiple methods and studies are required to demonstrate that parents' self-sorting is facilitated by resonance between race- and class-specific socialization (e.g., Bowles and Gintis 1976; Kohn 1969) and parenting styles (e.g., Baumrind 1971; Lareau 2011), on one hand, and the educational ideologies constructed by charter schools seeking to attract parents, on the other. Both distinct cultural skills (Lareau 2000, 2011) and positions in occupational hierarchies (Bowles and Gintis 1976; Kohn 1969) contribute to the reproduction of inequality through education (e.g., Bourdieu 1984; Weber 1946). To disentangle these influences on educational logics and enrollment outcomes, future studies should measure occupational positions or educational levels and cultural factors (e.g., vocabulary, interactional skills) to compare their impacts on school composition or parents' school choices.

Indeed, the present large-scale, observational study lays the groundwork for direct observational and longitudinal studies. For it is also possible that charter schools' expressed educational ideologies reflect manifest teaching practices, in which case they impart to different student populations different forms of socialization (Bowles and Gintis 1976) and consequently unequal access to dominant cultural and social capital (e.g., Golann 2015; Lareau 2011). Moreover, scholarship shows that the rise of formal discipline a means of social control (e.g., Durkheim 1961) and bodily regulation (e.g., Foucault 1977) is disproportionately applied to communities of color through strict disciplinary codes (e.g., "Zero-Tolerance" policies: Skiba 2000; Skiba and Rausch 2015; see also Morrill and Musheno 2018) and specialized programs (e.g., No Excuses; Golann 2015; Thernstrom and Thernstrom 2004). The relationships between educational ideologies and schools' disciplinary regimes are worthy of further study.

However, the web-based data of the present study require grounded observation through interviews or experiments to confirm the impacts of schools' heterogeneous identity claims (a process called "triangulation"; Powell et al. 2016), given that websites themselves may represent a form of "myth and ceremony" (Meyer and Rowan 1977). For unlike the present cross-section study, behavioral studies are capable of directly evidencing individual-level mechanisms—including the locally embedded decision-making processes by which charter schools purposefully frame their identities. Indeed, ethnographic studies are needed to capture how the pressures of accountability and market competition influence local processes of identity formation (e.g., see Appendix A). These forces could lead local actors to intentionally prioritize segregating marketing strategies for purposes of survival, resulting in the suspension of democratic values (e.g., diversity and inclusion) through "enactment" (Weick 1988) of economic logics. For example, qualitative analysis over time of meetings among charter school staff and boards—especially following episodes of administrative instability (e.g., Hallett 2010)—could reveal what interest groups push for identities that segregate social groups.

Similarly, with longitudinal data scholars could disentangle the causal influence of sociodemographics and educational ideology through parents' self-sorting into schools of choice (school selection effects), on one hand, and through schools' and CMOs' self-sorting into school districts (neighborhood selection effects), on the other.

Such insights have important implications for education policy, which has offered incentives and support for school choice programs including charter schools, vouchers, and home schooling since the federal No Child Left Behind law of 2001 (e.g., Hursh 2007; Lauen 2008), continuing with President Obama's Race to the Top initiative and Education Secretary Betsy DeVos (e.g., Brown 2016; Garvey 2017). At the state level, the promise of charter schools is "[providing] innovative learning opportunities and creative educational approaches to improve education" (Lubienski 2003; M. G. Assembly 2003)—a call that echoes through 90% of state charter policies (Wohlstetter, Smith, and Farrell 2013).

And in law, the U.S. Supreme Court has placed increasing legal limits on mandatory school district desegregation plans since *Brown v. Board of Education* (1954), which ruled that the legal doctrine of "separate but equal" or racially segregated schooling established under *Plessy v. Ferguson* (1896) violated the Equal Protection Clause of the Fourteenth Amendment. Specifically, the Supreme Court has prohibited court-ordered school district desegregation plans from integrating across district lines (namely, between urban and neighboring suburb districts; *Milliken v. Bradley* 1974), allowed for termination of said plans with only a good faith effort by the district (ensuring the end of segregation formally, but not in practice; *Board of Education v. Dowell* 1991), and forbade the use of racial classifications in student assignment plans to achieve integration (*Parents Involved in Community Schools v. Seattle School District No. 1* 2007). Such escalating legal barriers resulted in the dismissal of roughly half of states' desegregation plans by 2007 (Lutz 2011)—and in turn, gradual resegregation by race and income (Reardon et al. 2012; Reardon and Owens 2014). In fact, U.S. schools are overall more segregated now than they were before the *Brown* decision, with race and class often intersecting as African American and Hispanic students tend to be in the poorest schools (Orfield et al. 2014).

The upshot of this political and legal legacy is stratified school choice systems in which race and class groups congregate in schools with others like them (Malkus and Hatfield 2017; Wilson 2016). While students of color are isolated to strict disciplinary regimes meant to propel them to college (Golann 2015; Vasquez Heilig et al. 2019), relatively affluent, mobile white families concentrate (e.g., Holme 2002; Shrum et al. 1988) in schools and districts out of reach of the less privileged (e.g., Renzulli and Evans 2005; Saporito 2003). In such a system, schools of choice too are necessarily exclusionary: seeking to differentiate themselves within educational "markets" segregated by race and class (e.g., Fuller 2009; Lauen et al. 2015), such schools face pressure from authorizers and clients to cater to sociodemographic niches through location and admission decisions. This study (and those to follow) suggests a consequence of this organizational pressure: Charter schools select the best performers (e.g., Abdulkadiroglu et al. 2019; Adamson et al. 2015; Lacireno-Paquet et al. 2002) not universally, but among the race- and class-specific niche they seek to attract. Indeed, this research aims to document two segregating consequences of school choice policies: self-sorting by parents and schools and the niche-specific selection of students.

Many political and organizational solutions have been offered for segregated schools, including equitable access to information on school quality (e.g., Yettick 2016), accountability and fairness in charter schools' admissions procedures (e.g., Abdulkadiroglu et al. 2019; Holme 2002), or opening racially diverse schools (e.g., Roda and Wells 2013). Indeed, by illuminating an additional mechanism by which schools of choice participate in self-perpetuating inequalities, this work underscores the importance of incorporating into school choice programs stronger oversight of enrollment practices and encouragement for diverse schools.

Specifically, to obtain and reauthorize their charters, schools could be required to share their marketing strategies, student demographics, and evidence of attempts to recruit diverse student bodies. However, such remedies risk becoming ineffectual (or symbolic: Edelman 2016) insofar as legal enforcement is tied to organizational structures (e.g., public relations staff or protocols) rather than desirable outcomes (e.g., diverse student bodies). This is especially true given the many informal methods by which charter schools select or remove low-quality students (e.g., expulsion, intimidation; Adamson and Darling-Hammond 2016), which may undercut formal structures promoting integration. To mitigate this, financial incentives could be offered to charter schools that achieve greater diversity than their surrounding neighborhoods—a higher standard than raw enrollment figures and one more difficult to simulate. Nonetheless, such policies may depend on schools' and districts' freedom to consider students' race and class backgrounds in enrollment decisions, a legal bedrock the currently conservative U.S. Supreme Court continues to undermine (e.g., *Parents Involved in Community Schools v. Seattle School District No. 1* 2007). In such an environment, diversity-promoting programs may not only lack teeth in practice, but they also may be legally unenforceable. Reversing the trend of segregated schooling necessitates establishing legal precedents that defend and support integrating educational policies (Vasquez Heilig, Nelson, and Kronzer 2018).

So long as law and policy weaken organizational efforts toward integration, centrifugal social dynamics from organizational differentiation to neighborhood segregation to parents' racially charged perceptions of school quality will likely continue to pull students and families into charter schools where they meet fewer peers of race and class backgrounds different from their own—potentially undermining ethnic diversity and our civic ideals.

REFERENCES

Abdulkadiroglu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. 2019. "Do Parents Value School Effectiveness?" *American Economic Review* (forthcoming).

Abrahamson, Eric. 1996. "Management Fashion." *The Academy of Management Review* 21(1):254–85.

Achieve. 2010. *International Science Benchmarking Report. Taking the Lead in Science Education: Forging Next-Generation Science Standards*. Achieve, Inc.

Adamson, Frank, Bjorn Astrand, and Linda Darling-Hammond. 2016. *Global Education Reform: How Privatization and Public Investment Influence Education Outcomes*. New York & London: Routledge.

Adamson, Frank, Channa Cook-Harvey, and Linda Darling-Hammond. 2015. *Whose Choice? Student Experiences and Outcomes in the New Orleans School Marketplace*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

Adamson, Frank, and Linda Darling-Hammond. 2016. "The Critical Choice in American Education." Pp. 131–68 in *Global Education Reform: How Privatization and Public Investment Influence Education Outcomes*, edited by F. Adamson, B. Astrand, and L. Darling-Hammond. New York & London: Routledge.

Adamson, Frank, and Meredith Galloway. 2019. "Education Privatization in the United States: Increasing Saturation and Segregation." *Education Policy Analysis Archives* 27(0):129.

Airoldi, Edoardo M., and Jonathan M. Bischof. 2016. "Improving and Evaluating Topic Models and Other Models of Text." *Journal of the American Statistical Association* 111(516):1381–1403.

Aitchison, J., and S. M. Shen. 1980. "Logistic-Normal Distributions: Some Properties and Uses." *Biometrika* 67(2):261–72.

Albert, Stuart, and David A. Whetten. 1985. "Organizational Identity." *Research in Organizational Behavior* 7:263–95.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters. 2013. "Explaining Charter School Effectiveness." *American Economic Journal: Applied Economics* 5(4):1–27.

Anyon, Jean. 1980. "Social Class and the Hidden Curriculum of Work." *Journal of Education* 162(1):67–92.

Arora, Sanjeev, Rong Ge, Yonatan Halpern, David Mimno, Ankur Moitra, David Sontag, Yichen Wu, and Michael Zhu. 2013. "A Practical Algorithm for Topic Modeling with Provable Guarantees." Pp. 280–288 in *Proceedings of the 30th International Conference on Machine Learning*. Vol. 28. Atlanta: PMLR.

Arum, Richard. 1996. "Do Private Schools Force Public Schools to Compete?" *American Sociological Review* 61(1):29–46.

Arum, Richard. 2000. "Schools and Communities: Ecological and Institutional Dimensions." *Annual Review of Sociology* 26(1):395–418.

Asante, Molefi Kete, and Diane Ravitch. 1991. "Multiculturalism: An Exchange." *The American Scholar* 60(2):267–76.

Bail, Christopher A. 2014. "The Cultural Environment: Measuring Culture with Big Data." *Theory and Society* 43(3–4):465–82.

Bail, Christopher A., Taylor W. Brown, and Marcus Mann. 2017. "Channeling Hearts and Minds: Advocacy Organizations, Cognitive-Emotional Currents, and Public Conversation." *American Sociological Review* 82(6):1188–1213.

Baumrind, Diana. 1966. "Effects of Authoritative Parental Control on Child Behavior." *Child Development* 37(4):887–907.

Baumrind, Diana. 1971. "Current Patterns of Parental Authority." *Developmental Psychology* 4(1, Pt.2):1–103.

Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70(5):9–49.

Bennett, W. Lance, and Alexandra Segerberg. 2013. *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*. New York, NY: Cambridge University Press.

Benoit, Kenneth, Drew Conway, Benjamin E. Lauderdale, Michael Laver, and Slava Mikhaylov. 2016. "Crowd-Sourced Text Analysis: Reproducible and Agile Production of Political Data." *American Political Science Review* 110(2):278–95.

Berends, Mark. 2015. "Sociology and School Choice: What We Know After Two Decades of Charter Schools." *Annual Review of Sociology* 41(1):159–80.

Berger, Peter L., and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor Books.

Bischof, Jonathan M., and Edoardo M. Airoldi. 2012. "Summarizing Topical Content with Word Frequency and Exclusivity." Pp. 201–8 in *Proceedings of the 29th International Conference on Machine Learning*. New York: Omnipress.

Blei, David M. 2012. "Probabilistic Topic Models." *Communications of the ACM* 55(4):77–84.

Blei, David M., and John D. Lafferty. 2007. "A Correlated Topic Model of Science." *The Annals of Applied Statistics* 1(1):17–35.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3(Jan):993–1022.

Blinder, Alan. 2015. "Atlanta Educators Convicted in School Cheating Scandal." *The New York Times*, April 1.

Board of Education of Oklahoma City v. Dowell. 1991. 498 U.S. 237.

Bodine, Edward, Bruce Fuller, María-Fernanda González, Luis Huerta, Sandra Naughton, Sandra Park, and Laik Woon Teh. 2008. "Disparities in Charter School Resources—the Influence of State Policy and Community." *Journal of Education Policy* 23(1):1–33.

Bohr, Jeremiah, and Riley E. Dunlap. 2018. "Key Topics in Environmental Sociology, 1990–2014: Results from a Computational Text Analysis." *Environmental Sociology* 4(2):181–95.

Bourdieu, Pierre. 1977. *Outline of a Theory of Practice*. Cambridge University Press.

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press.

Bowles, Samuel, and Herbert Gintis. 1976. *Schooling in Capitalist America*. New York, NY: Basic Books.

Brewer, Marilynn B. 1991. "The Social Self: On Being the Same and Different at the Same Time." *Personality and Social Psychology Bulletin* 17(5):475–82.

Brint, Steven. 2013. "The 'Collective Mind' at Work: A Decade in the Life of U.S. Sociology of Education." *Sociology of Education* 86(4):273–79.

Brown, Emma. 2016. "School Choice Advocates Divided over Trump and His Education Pick, Betsy DeVos." *Washington Post*, December 12.

Brown v. Board of Education of Topeka. 1954. 347 U.S. 483.
Bruner, Jerome S. 1961. "The Act of Discovery." *Harvard Educational Review* 31:21–32.
Bryk, Anthony S., Penny Bender Sebring, Elaine Allensworth, John Q. Easton, and Stuart Luppescu. 2010. *Organizing Schools for Improvement: Lessons from Chicago*. University of Chicago Press.
Calarco, Jessica McCrory. 2011. "'I Need Help!' Social Class and Children's Help-Seeking in Elementary School." *American Sociological Review* 76(6):862–882.
Carnoy, Martin, Rebecca Jacobsen, Lawrence Mishel, and Richard Rothstein. 2005. *The Charter School Dust-Up*. Washington, D.C.: Economic Policy Institute.
Carroll, Glenn R. 1985. "Concentration and Specialization: Dynamics of Niche Width in Populations of Organizations." *American Journal of Sociology* 90:1262–83.
Carter, Samuel Casey. 2000. *No Excuses: Lessons from 21 High-Performing, High-Poverty Schools.* Washington, D.C.: Heritage Foundation.
Chait, Richard. 1979. "Mission Madness Strikes Our Colleges." *Chronicle of Higher Education* 18(36):36.
Chang, Jonathan, Jordan L. Boyd-Graber, Sean Gerrish, Chong Wang, and David M. Blei. 2009. "Reading Tea Leaves: How Humans Interpret Topic Models." *Advances in Neural Information Processing Systems* 288–96.
Chaves, Mark. 1996. "Ordaining Women: The Diffusion of an Organizational Innovation." *American Journal of Sociology* 101(4):840–73.
Chuang, Jason, Margaret E. Roberts, Brandon M. Stewart, Rebecca Weiss, Dustin Tingley, Justin Grimmer, and Jeffrey Heer. 2015. "TopicCheck: Interactive Alignment for Assessing Topic Model Stability." Pp. 175–184 in *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Denver, Colorado: Association for Computational Linguistics.
Chubb, John E., and Terry M. Moe. 1988. "Politics, Markets, and the Organization of Schools." *The American Political Science Review* 82(4):1066–87.
Chubb, John E., and Terry M. Moe. 1990. *Politics, Markets, and America's Schools*. Washington, D.C.: Brookings Institution Press.
Coburn, Cynthia E. 2004. "Beyond Decoupling: Rethinking the Relationship Between the Institutional Environment and the Classroom." *Sociology of Education* 77(3):211–44.
De Finetti, Bruno. 1975. *Theory of Probability: A Critical Introductory Treatment*. Vol. 1. London: John Wiley & Sons Ltd.
Dewey, John. 1897. *My Pedagogic Creed*. New York & Chicago: E.L. Kellogg & Company.
Dewey, John. 1938. *Experience and Education*. Indianapolis, IN: Kappa Delta Pi.
DiMaggio, Paul. 2015. "Adapting Computational Text Analysis to Social Science (and Vice Versa)." *Big Data & Society* 2(2):2053951715602908.
DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48(2):147–60.
DiMaggio, Paul J., and Walter W. Powell. 1991. "Introduction." Pp. 1–38 in *The New Institutionalism in Organizational Analysis*, edited by W. W. Powell and P. J. DiMaggio. Chicago: University of Chicago Press.
DiMaggio, Paul, Manish Nag, and David Blei. 2013. "Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding." *Poetics* 41(6):570–606.

Dreeben, Robert. 1968. *On What Is Learned in School*. Ontario: Addison Wesley.
Drudy, Sheelagh. 2008. "Gender Balance/Gender Bias: The Teaching Profession and the Impact of Feminisation." *Gender and Education* 20(4):309–23.
Durkheim, Emile. 1912. *The Elementary Forms of the Religious Life*. Courier Dover Publications.
Durkheim, Émile. 1961. *Moral Education*. Glencoe, New York: The Free Press.
Edelman, Lauren B. 1992. "Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law." *American Journal of Sociology* 97(6):1531–76.
Edelman, Lauren B. 2016. *Working Law: Courts, Corporations, and Symbolic Civil Rights*. University of Chicago Press.
Egami, Naoki, Christian J. Fong, Justin Grimmer, Margaret E. Roberts, and Brandon M. Stewart. 2018. "How to Make Causal Inferences Using Texts." Retrieved April 16, 2018 (https://arxiv.org/abs/1802.02163).
Eisenstein, Jacob, Amr Ahmed, and Eric P. Xing. 2011. "Sparse Additive Generative Models of Text." Pp. 1041–1048 in *Proceedings of the 28th International Conference on Machine Learning*, *ICML'11*. Bellevue, Washington, USA: Omnipress.
Elsbach, Kimberly D. 2003. "Organizational Perception Management." *Research in Organizational Behavior* 25:297–332.
Enns, Peter, Nathan Kelly, Jana Morgan, and Christopher Witko. 2016. "The Power of Economic Interests and the Congressional Economic Policy Agenda." *Scholars Strategic Network, June* 2.
Epple, Dennis, David Figlio, and Richard Romano. 2004. "Competition between Private and Public Schools: Testing Stratification and Pricing Predictions." *Journal of Public Economics* 88(7):1215–45.
Erickson, Heidi Holmes. 2017. "How Do Parents Choose Schools, and What Schools Do They Choose? A Literature Review of Private School Choice Programs in the United States." *Journal of School Choice* 11(4):491–506.
Evers, Williamson M. 2001. "Standards and Accountability." Pp. 205–237 in *A Primer on America's schools*, edited by T. M. Moe. Stanford, CA: Hoover Institution Press.
Fabricant, Michael, and Michelle Fine. 2012. *Charter Schools and the Corporate Makeover of Public Education: What's at Stake?* New York, NY: Teachers College Press.
Farrell, Caitlin, Priscilla Wohlstetter, and Joanna Smith. 2012. "Charter Management Organizations: An Emerging Approach to Scaling Up What Works." *Educational Policy* 26(4):499–532.
Fernandez, Manny. 2012. "El Paso Rattled by Scandal of 'Disappeared' Students." *The New York Times*, October 13.
Ferraro, Fabrizio, and Siobhán O'Mahony. 2012. "Managing the Boundaries of an 'Open' Project." Pp. 545–565 in *The Emergence of Organizations and Markets*, edited by J. F. Padgett and W. W. Powell. Princeton, NJ: Princeton University Press.
Finnigan, Kara S. 2007. "Charter School Autonomy: The Mismatch between Theory and Practice." *Educational Policy* 21(3):503–26.
Fligstein, N. 1990. *The Transformation of Corporate Control*. Harvard University Press.
Fligstein, Neil, Jonah Stuart Brundage, and Michael Schultz. 2017. "Seeing Like the Fed: Culture, Cognition, and Framing in the Failure to Anticipate the Financial Crisis of 2008." *American Sociological Review* 82(5):879–909.

Foucault, Michel. 1977. *Discipline and Punish: The Birth of the Prison*. New York, NY: Vintage Books.
Frankenberg, Erica. 2011. “Charter Schools: A Civil Rights Mirage?” *Kappa Delta Pi Record* 47(3):100–105.
Frankenberg, Erica, Jongyeon Ee, Jennifer B. Ayscue, and Gary Orfield. 2019. *Harming Our Common Future: America’s Segregated Schools 65 Years after Brown*. University of California, Los Angeles: Civil Rights Project/Proyecto Derechos Civiles.
Frankenberg, Erica, Stephen Kotok, Kai Schafft, and Bryan Mann. 2017. “Exploring School Choice and the Consequences for Student Racial Segregation within Pennsylvania’s Charter School Transfers.” *Education Policy Analysis Archives* 25(22).
Frankenberg, Erica, and Genevieve Siegel-Hawley. 2013. “A Segregating Choice?: An Overview of Charter School Policy, Enrollment Trends, and Segregation.” Pp. 129–44 in *Educational Delusions?: Why Choice Can Deepen Inequality and How to Make Schools Fair*. University of California Press.
Frankenberg, Erica, Genevieve Siegel-Hawley, Jia Wang, and Gary Orfield. 2010. *Choice Without Equity: Charter School Segregation and the Need for Civil Rights Standards*. University of California, Los Angeles: Civil Rights Project/Proyecto Derechos Civiles.
Friedland, Roger, and Robert Alford. 1991. “Bringing Society Back In: Symbols, Practices and Institutional Contradictions.” Pp. 232–63 in *The New Institutionalism in Organizational Analysis*, edited by W. Powell and P. Dimaggio. University of Chicago Press.
Friedman, Milton. 1955. “The Role of Government in Education.” Pp. 123–44 in *Economics and the public interest*, edited by R. A. Solo. New Brunswick, NJ: Rutgers University Press.
Fryer, Roland G. 2011. *Creating “No Excuses” (Traditional) Public Schools: Preliminary Evidence from an Experiment in Houston*. 17494. National Bureau of Economic Research, Working Paper.
Fuller, Bruce. 2009. “Policy and Place: Learning from Decentralized Reforms.” Pp. 855–875 in *Handbook of education policy research*, edited by G. Sykes, B. Schneider, and D. Plank. New York: Routledge.
Fuller, Bruce, Malena Arcidiacono, and Caitlin Kearns. 2020. “3. Civic Challengers and Organizational Variety.” P. 57 in *When Civic Activism Works – Pluralist Politics and School Reform in Los Angeles*. (forthcoming).
Furgeson, Joshua, Brian Gill, Joshua Haimson, Alexandra Killewald, Moira McCullough, Ira Nichols-Barrer, Bing-ru Teh, Natalya Verbitsky-Savitz, Melissa Bowen, Allison Demeritt, Paul Hill, and Robin Lake. 2012. *Charter-School Management Organizations: Diverse Strategies and Diverse Student Impacts*. Cambridge, MA: Mathematica Policy Research, Inc./Center on Reinventing Public Education.
Fusarelli, Lance D. 2004. “The Potential Impact of the No Child Left Behind Act on Equity and Diversity in American Education.” *Educational Policy* 18(1):71–94.
Garten, Justin, Joe Hoover, Kate M. Johnson, Reihane Boghrati, Carol Iskiwitch, and Morteza Dehghani. 2018. “Dictionaries and Distributions: Combining Expert Knowledge and Large Scale Textual Data Content Analysis.” *Behavior Research Methods* 50(1):344–61.
Garvey, Mike. 2017. “Betsy DeVos Refuses to Take a Stand Against Discrimination, Again.” *American Civil Liberties Union*. Retrieved January 10, 2018 (https://www.aclu.org/blog/lgbt-rights/lgbt-youth/betsy-devos-refuses-take-stand-against-discrimination-again).

Geiger, Atticus, Ignacio Cases, Lauri Karttunen, and Christopher Potts. 2019. "Posing Fair Generalization Tasks for Natural Language Inference - ACL Anthology." Pp. 4484–94 in. Hong Kong, China: Association for Computational Linguistics.

Gleason, Philip M. 2017. "What's the Secret Ingredient? Searching for Policies and Practices That Make Charter Schools Successful." *Journal of School Choice* 11(4):559–84.

Golann, Joanne W. 2015. "The Paradox of Success at a No-Excuses School." *Sociology of Education* 88(2):103–19.

Goldstein, Dana. 2017. "Obama Education Rules Are Swept Aside by Congress." *The New York Times*, March 9.

Goldstein, Harvey. 1987. "Multilevel Covariance Component Models." *Biometrika* 74(2):430–431.

Goldstein, Lisa S. 2008. "Kindergarten Teachers Making 'Street-Level' Education Policy in the Wake of No Child Left Behind." *Early Education and Development* 19(3):448–78.

Goodman, Joan F. 2013. "Charter Management Organizations and the Regulated Environment: Is It Worth the Price?" *Educational Researcher* 42(2):89–96.

Gouldner, Alvin W. 1954. *Patterns of Industrial Bureaucracy*. New York, NY: Free Press.

Graham, Jesse, Jonathan Haidt, and Brian A. Nosek. 2009. "Liberals and Conservatives Rely on Different Sets of Moral Foundations." *Journal of Personality and Social Psychology* 96(5):1029–46.

Grbic, Douglas, Frederic W. Hafferty, and Phillip K. Hafferty. 2013. "Medical School Mission Statements as Reflections of Institutional Identity and Educational Purpose: A Network Text Analysis." *Academic Medicine* 88(6):852–60.

Green, Erica L. 2017. "DeVos's Hard Line on New Education Law Surprises States." *The New York Times*, July 7.

Grimmer, Justin, and Gary King. 2011. "General Purpose Computer-Assisted Clustering and Conceptualization." *Proceedings of the National Academy of Sciences* 108(7):2643–50.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–97.

Guarino, Cassandra M., Lucrecia Santibañez, and Glenn A. Daley. 2006. "Teacher Recruitment and Retention: A Review of the Recent Empirical Literature." *Review of Educational Research* 76(2):173–208.

Hake, Richard R. 2008. "Language Ambiguities in Education Research." Retrieved September 17, 2018 (http://www.physics.indiana.edu/~hake).

Hallett, Tim. 2010. "The Myth Incarnate: Recoupling Processes, Turmoil, and Inhabited Institutions in an Urban Elementary School." *American Sociological Review* 75(1):52–74.

Hannan, Michael T., James N. Baron, Greta Hsu, and Ozgecan Kocak. 2006. "Organizational Identities and the Hazard of Change." *Industrial and Corporate Change, 15* 755–84.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2009. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." *Journal of Labor Economics* 27(3):349–83.

Hanushek, Eric A., John F. Kain, Steven G. Rivkin, and Gregory F. Branch. 2007. "Charter School Quality and Parental Decision Making with School Choice." *Journal of Public Economics* 91(5):823–48.

Harris, William Torrey. 1906. *The School City*. Syracuse: C.W. Bardeen.

Hartley, Matthew. 2003. "'There Is No Way without a Because': Revitalization of Purpose at Three Liberal Arts Colleges." *The Review of Higher Education* 27(1):75–102.

Hassel, Bryan C., and Paul E. Peterson. 2006. “Charter Schools: Mom and Pops or Corporate Design.” Pp. 149–160 in *Choice and competition in American education*. Lanham, MD: Rowman & Littlefield Publishers.

Haveman, Heather A., and Hayagreeva Rao. 1997. “Structuring a Theory of Moral Sentiments: Institutional and Organizational Coevolution in the Early Thrift Industry.” *American Journal of Sociology* 102:1606–51.

Heilig, Julian Vasquez. 2013. “Reframing the Refrain: Choice as a Civil Rights Issue.” *Texas Education Review* 1:83–94.

Heilig, Julian Vasquez, Jennifer Jellison Holme, Anthony V. LeClair, Lindsay D. Redd, and Derrick Ward. 2016. “Separate and Unequal: The Problematic Segregation of Special Populations in Charter Schools Relative to Traditional Public Schools.” *Stanford Law and Policy Review* 27:251–93.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. “The Weirdest People in the World?” *Behavioral and Brain Sciences* 33(2–3):61–83.

Hernández, Laura E. 2016. “Race and Racelessness in CMO Marketing: Exploring Charter Management Organizations’ Racial Construction and Its Implications.” *Peabody Journal of Education* 91(1):47–63.

Hirsch, Eric Donald. 1996. *The Schools We Need, and Why We Don’t Have Them*. New York, NY: Bantam Doubleday Dell.

Holme, Jennifer Jellison. 2002. “Buying Homes, Buying Schools: School Choice and the Social Construction of School Quality.” *Harvard Educational Review* 72(2):177–206.

Hsu, Greta. 2006. “Jacks of All Trades and Masters of None: Audiences’ Reactions to Spanning Genres in Feature Film Production.” *Administrative Science Quarterly* 51(3):420–50.

Huerta, Luis A. 2009. “Institutional v. Technical Environments: Reconciling the Goals of Decentralization in an Evolving Charter School Organization.” *Peabody Journal of Education* 84(2):244–61.

Huerta, Luis A., and Andrew Zuckerman. 2009. “An Institutional Theory Analysis of Charter Schools: Addressing Institutional Challenges to Scale.” *Peabody Journal of Education* 84(3):414–31.

Hursh, David. 2007. “Assessing ‘No Child Left Behind’ and the Rise of Neoliberal Education Policies.” *American Educational Research Journal* 44(3):493–518.

IST Research Corporation. 2017. *Scrapy Cluster*. Fredericksburg, VA.

Jha, Harsh K., and Christine M. Beckman. 2017. “A Patchwork of Identities: Emergence of Charter Schools as a New Organizational Form.” Pp. 69–107 in *Emergence*. Emerald Publishing Limited.

Johnson, Rucker C. 2019. *Children of the Dream: Why School Integration Works*. New York, NY: Basic Books.

Johnson, Terence J. 1972. *Professions and Power*. London: Macmillan.

Johnson, Victoria. 2007. “What Is Organizational Imprinting? Cultural Entrepreneurship in the Founding of the Paris Opera.” *American Journal of Sociology, 113* 97–127.

Jordan, Phyllis W., and Raegen Miller. 2017. *Who’s In: Chronic Absenteeism under the Every Student Succeeds Act*. Georgetown University, Washington, DC: FutureEd.

Jurafsky, Daniel, and James H. Martin. 2018. “19. Lexicons for Sentiment, Affect, and Connotation.” *Speech and Language Processing (3rd Edition)*. Retrieved April 1, 2019 (https://web.stanford.edu/~jurafsky/slp3/19.pdf).

Kaley, Anna, and Jakob Nielsen. 2019. "'About Us' Information on Corporate Websites." *Nielsen Norman Group*. Retrieved September 20, 2019 (https://www.nngroup.com/articles/about-us-information-on-websites/).

Kamenetz, Anya. 2015. *The Test*. New York, NY: PublicAffairs.

King, Brayden G., Elisabeth S. Clemens, and Melissa Fry. 2011. "Identity Realization and Organizational Forms: Differentiation and Consolidation of Identities Among Arizona's Charter Schools." *Organization Science* 22(3):554–72.

Kirschner, Paul A., John Sweller, and Richard E. Clark. 2006. "Why Minimal Guidance during Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching." *Educational Psychologist* 41(2):75–86.

Kohn, Alfie, Deborah Meier, and Tom Loveless. 2006. "Traditional, Progressive or a Bit of Both?" *The Washington Post*, May 9.

Kohn, Melvin. 1969. *Class and Conformity: A Study in Values*. Homewood, IL: The Dorsey Press.

Kornhaber, Mindy L. 2004. "Assessment, Standards, and Equity." *Handbook of Research on Multicultural Education* 2:91–109.

Kouloumpis, Efthymios, Theresa Wilson, and Johanna Moore. 2011. "Twitter Sentiment Analysis: The Good the Bad and the OMG!" in *Fifth International AAAI Conference on Weblogs and Social Media*.

Kozlowski, Austin C., Matt Taddy, and James A. Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings." *American Sociological Review* 84(5):905–49.

Krosnick, Jon A. 1999. "Survey Research." *Annual Review of Psychology* 50(1):537–67.

Kulkarni, Vivek, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2015. "Statistically Significant Detection of Linguistic Change." Pp. 625–635 in *Proceedings of the 24th International Conference on World Wide Web*, *WWW '15*. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Lacireno-Paquet, Natalie, Thomas T. Holyoke, Michele Moser, and Jeffrey R. Henig. 2002. "Creaming Versus Cropping: Charter School Enrollment Practices in Response to Market Incentives." *Educational Evaluation and Policy Analysis* 24(2):145–58.

Ladd, Helen F., ed. 1996. *Holding Schools Accountable: Performance-Based Reform in Education*. Washington, D.C.: Brookings Institution Press.

Lareau, Annette. 2000. *Home Advantage: Social Class and Parental Intervention in Elementary Education*. 2nd ed. Rowman & Littlefield Publishers.

Lareau, Annette. 2011. *Unequal Childhoods: Class, Race, and Family Life*. 2nd ed. Berkeley: University of California Press.

Lauen, Douglas Lee. 2008. "False Promises: The School Choice Provisions in the NCLB." Pp. 203–226 in *No Child Left Behind and the reduction of the achievement gap: Sociological perspectives on federal educational policy*.

Lauen, Douglas Lee, Bruce Fuller, and Luke Dauter. 2015. "Positioning Charter Schools in Los Angeles: Diversity of Form and Homogeneity of Effects." *American Journal of Education* 121(2):213–39.

Lazonder, Ard W., and Ruth Harmsen. 2016. "Meta-Analysis of Inquiry-Based Learning: Effects of Guidance." *Review of Educational Research* 86(3):681–718.

Leonardelli, Geoffrey J., Cynthia L. Pickett, and Marilynn B. Brewer. 2010. "Chapter 2 - Optimal Distinctiveness Theory: A Framework for Social Identity, Social Cognition, and Intergroup Relations." Pp. 63–113 in *Advances in Experimental Social Psychology*. Vol. 43, edited by M. P. Z. and J. M. Olson. Academic Press.

Levinson, Meira. 2012. *No Citizen Left Behind*. Harvard University Press.

Linn, Robert L. 2000. "Assessments and Accountability." *Educational Researcher* 29(2):4–16.

Lipsky, Michael. 1980. *Street-Level Bureaucracy: The Dilemmas of the Individual in Public Service*. New York: Russell Sage Foundation.

Loughran, Tim, and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66(1):35–65.

Louwerse, Max M. 2004. "Semantic Variation in Idiolect and Sociolect: Corpus Linguistic Evidence from Literary Texts." *Computers and the Humanities* 38(2):207–21.

Lubienski, Christopher. 2003. "Innovation in Education Markets: Theory and Evidence on the Impact of Competition and Choice in Charter Schools." *American Educational Research Journal* 40(2):395–443.

Lubienski, Christopher, Charisse Gulosino, and Peter Weitzel. 2009. "School Choice and Competitive Incentives: Mapping the Distribution of Educational Opportunities across Local Education Markets." *American Journal of Education* 115(4):601–47.

Lutz, Byron. 2011. "The End of Court-Ordered Desegregation." *American Economic Journal: Economic Policy* 3(2):130–68.

M. G. Assembly. 2003. "Maryland Public Charter School Act." *Title 9: Maryland Public Charter School Program*. Retrieved (http://www.marylandpublicschools.org/msde/programs/charter_schools/docs/md_charter_school_laws.htm).

Maaten, Laurens van der, and Geoffrey Hinton. 2008. "Visualizing Data Using T-SNE." *Journal of Machine Learning Research* 9(Nov):2579–2605.

Malkus, Nat. 2016. "Seeing Charters Differently: A New Approach to National Comparisons of Charter and Traditional Public Schools." *Journal of School Choice* 10(4):479–94.

Malkus, Nat, and Jenn Hatfield. 2017. *Differences by Design? Student Composition in Charter Schools with Different Academic Models*. Washington, D.C.: American Enterprise Institute (AEI).

March, James G., and Herbert A. Simon. 1958. *Organizations*. Cambridge, MA: Blackwell.

Massey, Douglas S., Jonathan Rothwell, and Thurston Domina. 2009. "The Changing Bases of Segregation in the United States." *The Annals of the American Academy of Political and Social Science* 626(1).

McNeil, Linda M. 2000. *Contradictions of School Reform: Educational Costs of Standardized Testing*. New York, NY: Routledge.

McShane, Michael Q., and Jenn Hatfield. 2015. *Measuring Diversity in Charter School Offerings*. Washington, D.C.: American Enterprise Institute (AEI).

Mehta, Jal. 2013. *The Allure of Order: High Hopes, Dashed Expectations, and the Troubled Quest to Remake American Schooling*. New York, NY: Oxford University Press.

Mehta, Jal. 2014. "When Professions Shape Politics: The Case of Accountability in K-12 and Higher Education." *Educational Policy* 28(6):881–915.

Meier, Deborah. 1995. *The Power of Their Ideas: Lessons for America from a Small School in Harlem.* Boston, MA: Beacon Press.

Meier, Deborah. 2002. *In Schools We Trust: Creating Communities of Learning in an Era of Testing and Standardization.* Boston, MA: Beacon Press.
Meier, Deborah. 2013a. "Explaining KIPP's 'SLANT.'" *Education Week - Bridging Differences*. Retrieved November 2, 2015 (http://blogs.edweek.org/edweek/Bridging-Differences/2013/04/slant_and_the_golden_rule.html?cmp=SOC-SHR-FB).
Meier, Deborah. 2013b. "Understanding 'No Excuses.'" *Education Week - Bridging Differences*. Retrieved November 2, 2015 (http://blogs.edweek.org/edweek/Bridging-Differences/2013/04/dear_elliott_were_going_to.html?cmp=SOC-SHR-FB).
Merton, Robert K. 1940. "Bureaucratic Structure and Personality." *Social Forces* 18:560–68.
Meyer, John W., John Boli, Thomas George M., and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103(1):144–81.
Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *The American Journal of Sociology* 83(2):340–63.
Meyer, John W., and Brian Rowan. 1978. "The Structure of Educational Organizations." Pp. 78–109 in *Environments and Organizations*. San Francisco, CA: Jossey-Bass.
Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." in *International Conference on Learning Representations Workshop*.
Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. 2013. "Distributed Representations of Words and Phrases and Their Compositionality." *Advances in Neural Information Processing Systems* 3111–3119.
Milliken v. Bradley. 1974. 418 U.S. 717.
Mimno, David, and Andrew McCallum. 2008. "Topic Models Conditioned on Arbitrary Features with Dirichlet-Multinomial Regression." Pp. 411–418 in *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence*, *UAI'08*. Helsinki, Finland: AUAI Press.
Mimno, David, Hanna M. Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. "Optimizing Semantic Coherence in Topic Models." Pp. 262–272 in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, *EMNLP '11*. Stroudsburg, PA, USA: Association for Computational Linguistics.
Miron, Gary, and Jessica Urschel. 2010. *Equal or Fair? A Study of Revenues and Expenditures in American Charter Schools*. Boulder, CO: National Education Policy Center.
Mohr, John W., and Petko Bogdanov. 2013. "Introduction—Topic Models: What They Are and Why They Matter." *Poetics* 41(6):545–69.
Mohr, John W., Robin Wagner-Pacifici, and Ronald L. Breiger. 2015. "Toward a Computational Hermeneutics." *Big Data & Society* 2(2):2053951715613809.
Monarrez, Tomas, Brian Kisida, and Matthew Chingos. 2019a. "Do Charter Schools Increase Segregation? First National Analysis Reveals a Modest Impact, Depending on Where You Look." *Education Next*, July 24.
Monarrez, Tomas, Brian Kisida, and Matthew Chingos. 2019b. *When Is a School Segregated? Making Sense of Segregation 65 Years after Brown v. Board of Education*. Urban Institute.
Moreno, Ivan. 2017. "US Charter Schools Put Growing Numbers in Racial Isolation." *AP News*. Retrieved January 10, 2018 (https://apnews.com/e9c25534dfd44851a5e56bd57454b4f5).
Morphew, Christopher C., and Matthew Hartley. 2006. "Mission Statements: A Thematic Analysis of Rhetoric Across Institutional Type." *The Journal of Higher Education* 77(3):456–71.

Morrill, Calvin, and Michael Musheno. 2018. *Navigating Conflict: How Youth Handle Trouble in a High-Poverty School*. University of Chicago Press.
NAACP. 2017. *NAACP Task Force on Quality Education Report*. Baltimore, MD: National Association for the Advancement of Colored People.
National Alliance for Public Charter Schools. 2018. "Data Dashboard." Retrieved October 1, 2018 (https://data.publiccharters.org/).
National Center for Education Statistics. 2018. "Digest of Education Statistics, 2017." *Enrollment and Percentage Distribution of Enrollment in Public Elementary and Secondary Schools, by Race/Ethnicity and Level of Education: Fall 1999 through Fall 2017*. Retrieved October 1, 2018 (https://nces.ed.gov/programs/digest/d17/tables/dt17_203.60.asp?).
National Center for Education Statistics. 2019. *Common Core of Data: Public Elementary/Secondary School Universe Survey*. Washington, D.C.: U.S. Department of Education. https://nces.ed.gov/ccd/pubschuniv.asp
Nelson, Laura K. 2017. "Computational Grounded Theory: A Methodological Framework." *Sociological Methods & Research* 1–40.
Nelson, Laura K., Derek Burk, Marcel Knudsen, and Leslie McCall. 2018. "The Future of Coding: A Comparison of Hand-Coding and Three Types of Computer-Assisted Text Analysis Methods." *Sociological Methods & Research* 0049124118769114.
Newmann, Fred M., BetsAnn Smith, Elaine Allensworth, and Anthony S. Bryk. 2001. "Instructional Program Coherence: What It Is and Why It Should Guide School Improvement Policy." *Educational Evaluation and Policy Analysis* 23(4):297–321.
Newsom, Walter, and C. Ray Hayes. 1991. "Are Mission Statements Worthwhile?" *Planning for Higher Education* 19(2):28–30.
Nielsen, Jakob, and John Morkes. 1997. "Concise, SCANNABLE, and Objective: How to Write for the Web." *Nielsen Norman Group*. Retrieved September 20, 2019 (https://www.nngroup.com/articles/concise-scannable-and-objective-how-to-write-for-the-web/).
Oakes, Jeannie, Martin Lipton, Lauren Anderson, and Jamy Stillman. 2013. *Teaching to Change the World*. 4th ed. Boulder, CO: Paradigm Publishers.
Orfield, Gary, and Jongyeon Ee. 2014. *Segregating California's Future: Inequality and Its Alternative 60 Years After Brown V. Board of Education*. University of California, Los Angeles: Civil Rights Project/Proyecto Derechos Civiles.
Orfield, Gary, Erica Frankenberg, Jongyeon Ee, and John Kuscera. 2014. *Brown at 60: Great Progress, a Long Retreat and an Uncertain Future*. University of California, Los Angeles: Civil Rights Project/Proyecto Derechos Civiles.
Orfield, Myron, Baris Gumus-Dawes, and Thomas Luce. 2013. "The State of Public Schools in Post-Katrina New Orleans: The Challenge of Creating Equal Opportunity." Pp. 159–84 in *Educational Delusions?: Why Choice Can Deepen Inequality and How to Make Schools Fair*. University of California Press.
Owens, Ann, Sean F. Reardon, and Christopher Jencks. 2016. "Income Segregation Between Schools and School Districts." *American Educational Research Journal* 53(4):1159–97.
Paino, Maria. 2018. "From Policies to Principals: Tiered Influences on School-Level Coupling." *Social Forces* 96(3):1119–54.
Paino, Maria, Rebecca L. Boylan, and Linda A. Renzulli. 2017. "The Closing Door: The Effect of Race on Charter School Closures." *Sociological Perspectives* 60(4):747–67.

Paino, Maria, Linda A. Renzulli, Rebecca L. Boylan, and Christen L. Bradley. 2014. "For Grades or Money? Charter School Failure in North Carolina." *Educational Administration Quarterly* 50(3):500–536.

Parents Involved in Community Schools v. Seattle School District No. 1. 2007. 551 U.S. 701.

Pedersen, Jesper Strandgaard, and Frank Dobbin. 2006. "In Search of Identity and Legitimation: Bridging Organizational Culture and Neoinstitutionalism." *American Behavioral Scientist* 49(7):897–907.

Pendergrass, Susan Aud, and Nora Kern. 2017. "The Case for Charters." Pp. 237–51 in *The Wiley Handbook of School Choice*, edited by R. A. Fox and N. K. Buchanan. Hoboken, NJ: John Wiley & Sons, Inc.

Pennebaker, James W., Martha E. Francis, and Roger J. Booth. 2001. "Linguistic Inquiry and Word Count: LIWC 2001." *Mahway: Lawrence Erlbaum Associates* 71.

Petrilli, Michael. 2012. *The Diverse Schools Dilemma: A Parent's Guide to Socioeconomically Mixed Public Schools*. Dayton, Ohio: Thomas B. Fordham Institute.

Pettigrew, Andrew M. 1979. "On Studying Organizational Cultures." *Administrative Science Quarterly* 24(4):570–81.

Plessy v. Ferguson. 1896. 163 U.S. 537.

Porac, Joseph F., Howard Thomas, Fiona Wilson, Douglas Paton, and Alaina Kanfer. 1995. "Rivalry and the Industry Model of Scottish Knitwear Producers." *Administrative Science Quarterly* 203–227.

Porter, Martin F. 1980. "An Algorithm for Suffix Stripping." *Program* 14(3):130–137.

Posey-Maddox, Linn, Shelley McDonough Kimelberg, and Maia Cucchiara. 2014. "Middle-Class Parents and Urban Public Schools: Current Research and Future Directions." *Sociology Compass* 8(4):446–56.

Potter, Halley, and Kimberly Quick. 2018. *Diverse-by-Design Charter Schools*. New York, NY: The Century Foundation.

Powell, Walter W., Aaron Horvath, and Christof Brandtner. 2016. "Click and Mortar: Organizations on the Web." *Research in Organizational Behavior* 36:101–120.

Preston, Courtney, Ellen Goldring, Mark Berends, and Marisa Cannata. 2012. "School Innovation in District Context: Comparing Traditional Public Schools and Charter Schools." *Economics of Education Review* 31(2):318–30.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramirez, Francisco O., and John W. Meyer. 2012. "Toward Post-National Societies and Global Citizenship." *Multicultural Education Review* 4(1):1–28.

Rao, Hayagreeva, Philippe Monin, and Rodolphe Durand. 2003. "Institutional Change in Toque Ville: Nouvelle Cuisine as an Identity Movement in French Gastronomy." *American Journal of Sociology* 108(4):795–843.

Ravitch, Diane. 2010. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books.

Raymond, Margaret E., Edward Cremata, Devora Davis, Kathleen Dickey, Kristina Lawyer, Yohannes Negassi, and James L. Woodworth. 2013. *National Charter School Study*. Stanford, CA: Center for Research on Education Outcomes (CREDO), Stanford University.

Reardon, Sean F., Elena Tej Grewal, Demetra Kalogrides, and Erica Greenberg. 2012. "Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools." *Journal of Policy Analysis and Management* 31(4):876–904.

Reardon, Sean F., Demetra Kalogrides, and Ken Shores. 2017. *The Geography of Racial/Ethnic Test Score Gaps*. *Working Paper*. 16–10. Stanford, CA: Center for Education Policy Analysis.

Reardon, Sean F., and Ann Owens. 2014. "60 Years After Brown: Trends and Consequences of School Segregation." *Annual Review of Sociology* 40(1):199–218.

Reardon, Sean F., Joseph Townsend, and Lindsay Fox. 2017. "A Continuous Measure of the Joint Distribution of Race and Income Among Neighborhoods." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3(2):34–62.

Reay, Diane, Gill Crozier, David James, Sumi Hollingworth, Katya Williams, Fiona Jamieson, and Phoebe Beedell. 2008. "Re-Invigorating Democracy?: White Middle Class Identities and Comprehensive Schooling." *The Sociological Review* 56(2):238–55.

Řehůřek, Radim, and Petr Sojka. 2010. "Software Framework for Topic Modelling with Large Corpora." Pp. 45–50 in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*.

Renzulli, Linda. 2014. "Educational Transformations and Why Sociology Should Care." *Social Currents* 1(2):149–56.

Renzulli, Linda A., Ashley B. Barr, and Maria Paino. 2015. "Innovative Education? A Test of Specialist Mimicry or Generalist Assimilation in Trends in Charter School Specialization Over Time." *Sociology of Education* 88(1):83–102.

Renzulli, Linda A., and Lorraine Evans. 2005. "School Choice, Charter Schools, and White Flight." *Social Problems* 52(3):398–418.

Richardson, Leonard. 2007. *Beautiful Soup Documentation*. New York, NY.

Ritter, Gary, Nathan Jensen, Brian Kisida, and Joshua McGee. 2010. "A Closer Look at Charter Schools and Segregation." *Education Next*, April 27.

Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airoldi. 2016. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association* 111(515):988–1003.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2016. "Navigating the Local Modes of Big Data." in *Computational Social Science: Discovery and Prediction*, edited by A. R. Michael. New York: Cambridge University Press.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2019. "Stm: R Package for Structural Topic Models." *Journal of Statistical Software* 91(2):1–40.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, and Edoardo M. Airoldi. 2013. "The Structural Topic Model and Applied Social Science." Pp. 1–20 in *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58(4):1064–82.

Roda, Allison, and Amy Stuart Wells. 2013. "School Choice Policies and Racial Segregation: Where White Parents' Good Intentions, Anxiety, and Privilege Collide." *American Journal of Education* 119(2):261–93.

Roth, Erin, Abel McDaniels, Catherine Brown, and Neil Campbell. 2017. "The Progressive Case for Charter Schools." *Center for American Progress*. Retrieved

(https://www.americanprogress.org/issues/education-k-12/news/2017/10/24/440833/the-progressive-case-for-charter-schools/).
Rothstein, Jesse M. 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *American Economic Review* 96(4):1333–50.
Rubin, Donald B. 1976. "Inference and Missing Data." *Biometrika* 63(3):581–592.
Salton, G., A. Wong, and C. S. Yang. 1975. "A Vector Space Model for Automatic Indexing." *Communications of the ACM* 18(11):613–620.
Salton, Gerard, ed. 1971. *The SMART Retrieval System: Experiments in Automatic Document Processing*. Englewood Cliffs, NJ: Prentice-Hall.
Salton, Gerard, and Christopher Buckley. 1988. "Term-Weighting Approaches in Automatic Text Retrieval." *Information Processing & Management* 24(5):513–523.
Sapir, Edward. 1949. *Culture, Language and Personality*. Berkeley, CA: University of California Press.
Saporito, Salvatore. 2003. "Private Choices, Public Consequences: Magnet School Choice and Segregation by Race and Poverty." *Social Problems* 50(2):181–203.
Sawchuk, Andrew Ujifusa, Alyson Klein, Stephen. 2017. "A Guide to State ESSA Plans: Goals, Teacher Quality, and More." *Education Week*. Retrieved January 8, 2018 (https://www.edweek.org/ew/section/multimedia/a-guide-to-state-essa-plans-goals-teacher-quality.html).
Schakel, Adriaan M. J., and Benjamin J. Wilson. 2015. "Measuring Word Significance Using Distributed Representations of Words." *ArXiv:1508.02297 [Cs]*.
Schneider, Barbara, Martin Carnoy, Jeremy Kilpatrick, William H. Schmidt, and Richard J. Shavelson. 2007. *Estimating Causal Effects Using Experimental and Observational Designs. Think Tank White Paper*. Washington, D.C.: American Educational Research Association.
Schwartz, H. Andrew, and Lyle H. Ungar. 2015. "Data-Driven Content Analysis of Social Media: A Systematic Overview of Automated Methods." *The ANNALS of the American Academy of Political and Social Science* 659(1):78–94.
Scott, Janelle. 2018. "The Problem We All Still Live With: Neo-Plessyism, and School Choice Policies in the Post-Obama Era." in *Choosing Charters: Better Schools or More Segregation?*, edited by I. C. Rotberg and J. L. Glazer. New York, NY: Teachers College Press.
Scott, Janelle T. 2013. "A Rosa Parks Moment? School Choice and the Marketization of Civil Rights." *Critical Studies in Education* 54(1):5–18.
Scott, W. Richard. 2001. *Institutions and Organizations, 2nd Ed*. Thousand Oaks, CA: Sage.
Scott, W. Richard, and John W. Meyer. 1983. "The Organization of Societal Sectors." Pp. 129–54 in *Organizational Environments: Ritual and Rationality*, edited by J. W. Meyer and W. R. Scott. Newbury Park, CA: Sage.
Selznick, Philip. 1949. *TVA and the Grass Roots: A Study in the Sociology of Formal Organization*. Berkeley and Los Angeles: University of California Press.
Selznick, Philip. 1957. *Leadership in Administration: A Sociological Interpretation*. University of California Press.
Shane, Janelle. 2019. *You Look Like a Thing and I Love You: How Artificial Intelligence Works and Why It's Making the World a Weirder Place*. Little, Brown.
Shanker, Albert. 1988. "Restructuring Our Schools." *Peabody Journal of Education* 65(3):88–100.

Shrum, Wesley, Neil H. Cheek, and Saundra MacD. Hunter. 1988. "Friendship in School: Gender and Racial Homophily." *Sociology of Education* 61(4):227–39.
Simon, Herbert A. 1946. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. 3rd ed. New York, NY: Free Press.
Sivak, Elizaveta, and Ivan Smirnov. 2019. "Parents Mention Sons More Often than Daughters on Social Media." *Proceedings of the National Academy of Sciences* 116(6):2039–41.
Skiba, Russell J. 2000. *Zero Tolerance, Zero Evidence: An Analysis of School Disciplinary Practice*. Indiana University, Bloomington: Education Policy Center.
Skiba, Russell J., and M. Karega Rausch. 2015. "Reconsidering Exclusionary Discipline: The Efficacy and Equity of out-of-School Suspension and Expulsion." Pp. 116–38 in *Handbook of Classroom Management*, edited by E. T. Emmer and E. J. Sabornie. New York: Routledge.
Sloan, Kris. 2006. "Teacher Identity and Agency in School Worlds: Beyond the All-Good/All-Bad Discourse on Accountability-Explicit Curriculum Policies." *Curriculum Inquiry* 36(2):119–52.
Sontag, David, and Daniel M. Roy. 2009. "Complexity of Inference in Topic Models." in *Advances in Neural Information Processing: Workshop on Applications for Topic Models: Text and Beyond*.
Spillane, J. P., and Patricia Burch. 2006. "The Institutional Environment and Instructional Practice: Changing Patterns of Guidance and Control in Public Education." Pp. 87–102 in *The new institutionalism in education*, edited by H.-D. Meyer and B. Rowan.
Spirling, Arthur, and Pedro L. Rodriguez. 2019. "Word Embeddings: What Works, What Doesn't, and How to Tell the Difference for Applied Research."
StataCorp. 2017a. *Stata 15 Multiple Imputation Reference Manual*. College Station, TX: Stata Press.
StataCorp. 2017b. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.
Steffe, Leslie P., and Jerry Edward Gale. 1995. *Constructivism in Education*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Stinchcombe, Arthur L. 1965. "Social Structure and Organizations." Pp. 142–93 in *Handbook of Organizations*, edited by J. March. Chicago: Rand-McNally.
Stone, Philip J., Dexter C. Dunphy, and Marshall S. Smith. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. Oxford, England: M.I.T. Press.
Strathern, Marilyn. 2000. *Audit Cultures: Anthropological Studies in Accountability, Ethics, and the Academy*. Psychology Press.
Swaminathan, Anand. 2001. "Resource Partitioning and the Evolution of Specialist Organizations: The Role of Location and Identity in US Wine." *Academy of Management Journal* 44:1169–85.
Swidler, Ann. 1979. *Organization without Authority: Dilemmas of Social Control in Free Schools*. Harvard University Press.
TensorFlow. 2018. "Vector Representations of Words." *TensorFlow*. Retrieved September 21, 2018 (https://www.tensorflow.org/tutorials/representation/word2vec).
Teske, Paul, Jody Fitzpatrick, and Gabriel Kaplan. 2006. "The Information Gap?" *Review of Policy Research* 23(5):969–81.
Thernstrom, Abigail, and Stephan Thernstrom. 2004. *No Excuses: Closing the Racial Gap in Learning*. Simon and Schuster.

Tobin, Joseph Jay, David Yen Ho Wu, and Dana H. Davidson. 1991. *Preschool in Three Cultures: Japan, China, and the United States*. Yale University Press.

Tolbert, Pamela S., and Lynne G. Zucker. 1983. “Institutional Sources of Change in the Formal Structure of Organizations: The Diffusion of Civil Service Reform, 1880-1935.” *Administrative Science Quarterly* 28(1):22–39.

Towns, John, Timothy Cockerill, Maytal Dahan, Ian Foster, Kelly Gaither, Andrew Grimshaw, Victor Hazlewood, Scott Lathrop, Dave Lifka, Gregory D. Peterson, Ralph Roskies, J. Ray Scott, and Nancy Wilkins-Diehr. 2014. “XSEDE: Accelerating Scientific Discovery.” *Computing in Science & Engineering* 16(5):62–74.

Turian, Joseph, Lev-Arie Ratinov, and Yoshua Bengio. 2010. “Word Representations: A Simple and General Method for Semi-Supervised Learning.” Pp. 384–394 in *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Uppsala, Sweden: Association for Computational Linguistics.

Tuttle, Christina Clark, Bing-ru Teh, Ira Nichols-Barrer, Brian P. Gill, and Philip Gleason. 2010. *Student Characteristics and Achievement in 22 KIPP Middle Schools. Final Report*. Washington, D.C.: Mathematica Policy Research, Inc.

U.S. Census Bureau. 2018. *American Community Survey, 2012-16 Summary File Data*. Washington, D.C.

U.S. Department of Education. 2018. “EDFacts Data Files.” Retrieved September 15, 2016 (http://www2.ed.gov/about/inits/ed/edfacts/data-files/index.html).

Van Rossum, Guido, and Fred L. Drake. 2011. *Python Language Reference Manual*. United Kingdom: Network Theory Limited.

Vasquez Heilig, Julian, T. Jameson Brewer, and Yohuru Williams. 2019. “Choice without Inclusion?: Comparing the Intensity of Racial Segregation in Charters and Public Schools at the Local, State and National Levels.” *Education Sciences* 9(3):205–21.

Vasquez Heilig, Julian, Steven Nelson, and Matt Kronzer. 2018. “Does the African American Need Separate Charter Schools?” *Law & Inequality: A Journal of Theory and Practice* 36(2).

Waite, Chelsea. 2019. *A View from the Canopy: Building Collective Knowledge on School Innovation*. Christensen Institute.

Wang, Ke, Amy Rathbun, and Lauren Musu. 2019. *School Choice in the United States: 2019*. Washington, D.C.: National Center for Education Statistics.

Watanabe, Maika. 2007. “Displaced Teacher and State Priorities in a High-Stakes Accountability Context.” *Educational Policy* 21(2):311–68.

Watson, David, and Lee Anna Clark. 1999. *The PANAS-X: Manual for the Positive and Negative Affect Schedule-Expanded Form*. University of Iowa.

Weber, Max. 1905. *The Protestant Ethic and the Spirit of Capitalism*. Mineola, N.Y.: Dover Publications, Inc.

Weber, Max. 1946. “Class, Status, Party.” Pp. 180–95 in *From Max Weber: Essays in Sociology*. New York, NY: Oxford University Press.

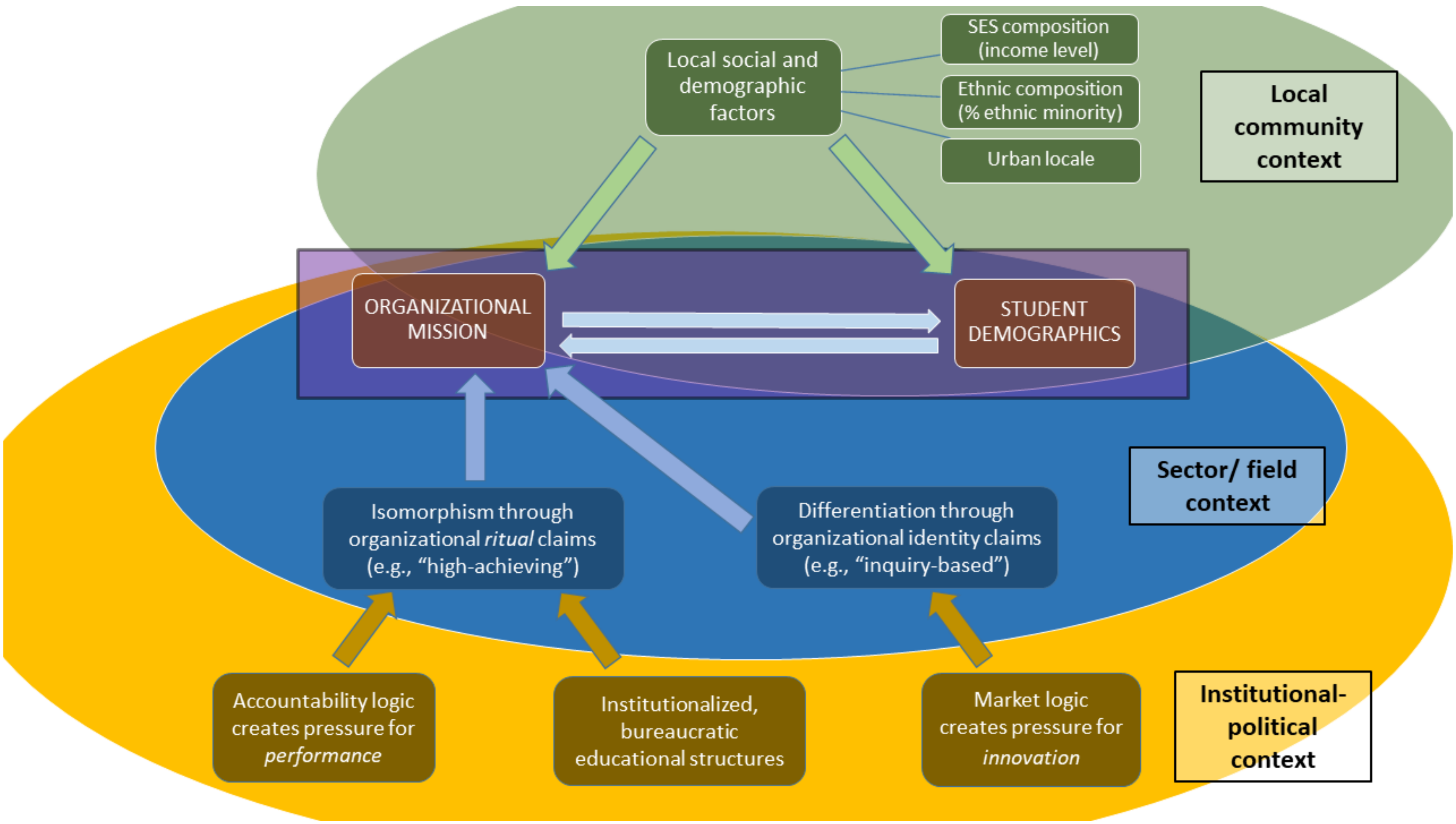
Weber, Max. 1968. *Economy and Society: An Outline of Interpretive Sociology*. Vol. 1. edited by G. Roth and C. Wittich. Berkeley: University of California Press.

Weick, Karl E. 1976. “Educational Organizations as Loosely Coupled Systems.” *Administrative Science Quarterly* 21(1):1–19.

Weick, Karl E. 1988. “Enacted Sensemaking in Crisis Situations.” *Journal of Management Studies* 25(4):305–17.

Westphal, James D., Ranjay Gulati, and Stephen M. Shortell. 1997. "Customization or Conformity? An Institutional and Network Perspective on the Content and Consequences of TQM Adoption." *Administrative Science Quarterly* 42(2):366–94.
Westphal, James D., and Edward J. Zajac. 1994. "Substance and Symbolism in CEOs' Long-Term Incentive Plans." *Administrative Science Quarterly* 39(3):367–90.
Whetten, David A. 2006. "Albert and Whetten Revisited: Strengthening the Concept of Organizational Identity." *Journal of Management Inquiry* 15(3):219–34.
Whorf, Benjamin Lee. 1940. "Science and Linguistics." *Technology Review* 42(6):229–31, 247–48.
Williams, Jaime, William Smythe, Thomas Hadjistavropoulos, David C. Malloy, and Ronald Martin. 2005. "A Study of Thematic Content in Hospital Mission Statements: A Question of Values." *Health Care Management Review* 30(4):304–14.
Wilson, Benjamin J., and Adriaan M. J. Schakel. 2015. "Controlled Experiments for Word Embeddings." *ArXiv:1510.02675 [Cs]*.
Wilson, Terri S. 2016. "Contesting the Public School: Reconsidering Charter Schools as Counterpublics." *American Educational Research Journal* 53(4):919–52.
Wohlstetter, Priscilla, Joanna Smith, and Caitlin C. Farrell. 2013. *Choices and Challenges: Charter School Performance in Perspective*. Cambridge, MA: Harvard Education Press.
Woodworth, James L., Margaret E. Raymond, Chunping Han, Negasi Yohannes, Richardson W. Payton, and Will Snow. 2017. *Charter Management Organizations*. Stanford, CA: Center for Research on Education Outcomes, Stanford University.
Yettick, Holly. 2016. "Information Is Bliss: Information Use by School Choice Participants in Denver." *Urban Education* 51(8):859–90.
Zeehandelaar, Dara, and Amber M. Winkler. 2013. *What Parents Want: Education Preferences and Trade-Offs*. Dayton, Ohio: Thomas B. Fordham Institute.
Zhang, Yahong, and Kaifeng Yang. 2008. "What Drives Charter School Diffusion at the Local Level: Educational Needs or Political and Institutional Forces?" *Policy Studies Journal* 36(4):571–91.
Zucker, Lynne G. 1977. "The Role of Institutionalization in Cultural Persistence." *American Sociological Review* 42(5):726–43.
Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Legitimacy Discount." *American Journal of Sociology* 104:1398–1438.

APPENDIX A: INSTITUTIONAL CONTEXTS SHAPING CHARTER SCHOOLS' IDEOLOGIES



APPENDIX B: TECHNICAL NOTES ON METHODS

Structural Topic Models

STM has several critical differences with LDA (Roberts et al. 2014:1067; Roberts, Stewart, and Airoldi 2016). Most importantly, STM allows each document to have its own prior distribution over topics by incorporating document metadata, *X*, enabling documents' "topical prevalence"—*what* topics they talk about—to vary with observed characteristics. For instance, a school in an affluent neighborhood may in its website talk more about course offerings than do schools elsewhere. Moreover, building off an improvement on LDA called the Correlated Topic Model (Blei and Lafferty 2007) and prior efforts at incorporating covariates in topic models (Mimno and McCallum 2008), STM allows the topic proportions (θ) to be correlated by replacing the Dirichlet with a logistic-normal linear model (Aitchison and Shen 1980). This more flexible approach replaces the global priors (α, β_k) in earlier approaches (e.g., Blei and Lafferty 2007; Blei et al. 2003) with covariate-specific priors and global topic covariance (Roberts et al. 2013). Thus, topic proportions are distributed over a logistic-normal distribution, with topical prevalence coefficients γ and topic covariance matrix ∑:

$$\theta_d \sim \text{LogisticNormal}(X_{d\gamma}, \textstyle\sum)$$

Another key difference with LDA is that with STM, the topic-word matrix or "topical content"—*how* they talk about topics—varies with a second set of covariates, *Y*. For instance, a school in a poor neighborhood may in its online coverage of courses focus more on technical ones than those in the humanities. Building on prior efforts to vary topical content with covariates (Eisenstein, Ahmed, and Xing 2011), STM accomplishes this through a multinomial logit with three effects dependent on content covariates (y_d)—topics ($\kappa_k^{(t)}$), covariates ($\kappa_{y_d}^{(c)}$), and topic-covariate interactions ($\kappa_{y_d,k}^{(i)}$)—responsible for sparse deviations from baseline word y_d frequencies (*m*). Thus, the document-specific distribution of words over each topic ($\beta_{d,k}$) is formed using an exponential of *m* and these three vectors (each the length of the vocabulary) as follows:

$$\beta_{d,k} \propto \exp(m + \kappa_k^{(topic)} + \kappa_{y_d}^{(cov)} + \kappa_{y_d,k}^{(int)})$$

I did not include any topical content covariates in my models, as attempts to do so were computationally intractable and their models failed to converge.

To enhance interpretability, STM uses a Laplace prior to induce sparsity for the κ parameters (i.e., sets many of them to zero). This is essentially an application of L1 regularization[19] to each update of the model using document-level information. On the other hand, a zero-mean Gaussian distribution is used as prior for the topic prevalence parameters (γ and ∑), which prevents

[19] Also known as lasso regression, L1 regularization involves adding a constraint during the minimization of the loss function. This provides a solution to overfitting, i.e., model estimates lacking generalizability due to over-reliance on training data. While overfitting is less an issue in simple linear regression models, deep learning models (i.e., analysis of "big data") are more complex and prone to overfitting.

overfitting by shrinking coefficients toward zero (without inducing sparsity). Finally, the model estimates are initialized using spectral decomposition (i.e., non-negative matrix factorization), which produces more consistent results than random or LDA initialization (Arora et al. 2013; Roberts et al. 2019; Roberts, Stewart, and Tingley 2016).

While model initialization, priors, and many other parameters are configurable in STM, modeling my relatively simple, moderately sized corpus doesn't call for minute changes. As such, I use the default settings for STM except where described above.

Word embeddings

While a few alternative model architectures have been proposed for neural-net word embedding models, in this paper I use the "continuous skip-gram" model, which is the most accurate option for semantic comparisons (Mikolov, Chen, et al. 2013)—the focus of this paper—and is best suited for large data sets (TensorFlow 2018) like my charter schools database. Continuous skip-gram seeks to classify a word given each other word in its context, allowing each word vector to have a separate impact on its neighbors.

In my implementation of word embeddings, I take advantage of several extensions to the original continuous skip-gram model: noise reduction, undersampling of frequent words, and detection of common phrases in the corpus (Mikolov, Sutskever, et al. 2013). Common phrases, also called multi-word expressions, often possess unique meanings compared to their individual words; for instance, compare the meaning of the phrase "special education" to the individual terms "special" and "education". My model accepts phrases so long as their 'collocation score'—the ratio of a given word pair's collocations divided by the product of each word's individual appearances—exceeds some threshold. Mathematically, this score is derived as follows:

$$\mathrm{score}(w_i, w_j) = \frac{\mathrm{count}(w_i w_j) - \delta}{\mathrm{count}(w_i) \times \mathrm{count}(w_j)} \tag{3}$$

where w_i and w_j are two words in the corpus, $w_i w_j$ is their collocation, and δ is a "discounting coefficient" that makes infrequent phrases less likely (Mikolov, Sutskever, et al. 2013:6).

Computationally, word embeddings learn word vectors using a 'neural network', a machine learning architecture that tries to predict (in this case) the connections between words using a sample of observations of word connections or 'training words' (e.g., Turian, Ratinov, and Bengio 2010). In technical terms, word embeddings project the semantic connections observed in the sample (using initially random weights) through a hidden layer and into an output matrix, the error of which is used to update and optimize the weights over several iterations. The size of the hidden layer equals *k*, the number of vector space dimensions in the model.

Additionally, in calculating similarities, word2vec normalizes all vectors to a length of 1, essentially projecting all word vectors along the surface of a high-dimensional sphere. Unit normalization improves computational efficiency (Wilson and Schakel 2015), but it may lose some information on the consistency with which words appear in their contexts (Schakel and Wilson 2015).

For purposes of efficiency, I use the `gensim` defaults for some parameters: 10 iterations through the data, 5 noise words (for noise reduction), an initial learning rate of 0.025, and a discounting coefficient of 3. To improve the model's ability to capture semantic nuances—

without inducing unnecessary computational burden—I expand the defaults for a few other parameters: word context windows of size eight (to better capture syntactic relations: Mikolov, Chen, et al. 2013; Spirling and Rodriguez 2019), 20000 words per sample, a hidden layer of size 300, and a phrase detection threshold of 8.0 (to better capture candidate multi-word expressions). As suggested in Mikolov et al. (2013:4), I use a negative sampling exponent of 0.75 to subsample frequent words.

Finally, while the number of dimensions, k, in the word embedding vector space is high (typically 100-300), this is nonetheless far smaller than the number of words in the corpus. Indeed, the dimensionality is low compared to the older and less efficient (though still common) term-document matrix: words are represented as rows, documents as columns, and each entry as the frequency or weight of a given term in a given document (Salton, Wong, and Yang 1975). For a vocabulary of size N and corpus of D documents, word embeddings represent all words and their associations using N x k dimensions, while the document-term matrix represents words using N x D dimensions. Even for a relatively small corpus like 6,300 websites, the latter matrix is enormously larger; consequently, working with the former matrix requires significantly less time and computational resources.

APPENDIX C: EXAMPLES OF CHARTER SCHOOL WEBSITES

LOW in inquiry-based learning (IBL) emphasis: Phoenix College Preparatory Academy, Phoenix, AZ (https://www.phoenixcollege.edu/pc-prep-academy)

High school and college in one...

The Phoenix College Preparatory Academy (PCPA) is a charter high school accredited by the North Central Association Commission on Accreditation and School Improvement (NCA CASI), an accreditation division of AdvancED.

Students at our academy attend classes on the beautiful campus of Phoenix College, and in addition to the resources of the high school, students have access to the colleges' computer labs, libraries and other services and facilities. PCPA students have the opportunity to interact with community college students who serve as both mentors and tutors.

PCPA maintains the highest academic standards to ensure students meet all state requirements for a high school diploma. It is also possible that by graduation students can complete several community college courses.

IT'S LIKE RECEIVING A PRIVATE EDUCATION FOR FREE....

As a public charter high school, Phoenix College Preparatory Academy adheres to the open enrollment policies prescribed by the Arizona Department of Education. PCPA does not charge tuition for high school classes. Students in good academic standing may qualify for free college tuition as well.

Why attend our academy?

- Individualized instruction from state certified and/or highly qualified teachers in a small class setting
- Obtain your high school diploma
- Join in community college extra-curricular activities & organizations
- Access to college resources and facilities, including free tutoring
- May qualify to receive free tuition for college courses at any of the Maricopa Community Colleges

Phoenix College Preparatory Academy is a free public high school and does not require any citizenship or immigration status information or documentation to enroll students into high school classes. Students are not required to take college courses and are enrolled on a first come, first served, basis.

Mission Statement

Through a shared vision, Phoenix College Preparatory Academy, supported by Phoenix College, is committed to creating and sustaining a community where all learners will pursue high standards to succeed in college and career

Vision

- Every PCPA student will complete the general high school requirements in order to be admitted to a four year postsecondary institution
- Each student will have the opportunity to earn at least 30 college credits, or complete one year of college credit requirements
- 20% of students will complete their Associate's degree prior to high school graduation

---

HIGH in IBL emphasis: Anne Frank Inspire Academy, San Antonio, TX (https://www.braination.net/Anne-Frank-Inspire-Academy)

A Preeminent 21st Century School K–12

How do you make a free, public, 21st century school that is caring, creative, fun, innovative, and sustainable? How do you ensure that school retains a rigorous curriculum that pushes every student to be and do his or her best? You create an Inspire Academy!

The curriculum at our Inspire Academies adheres to state standards in all subjects, including the core areas of language arts, mathematics, science, and social studies. But in order to fully address student needs, we have designed a "three tiered curriculum" that not only meets state standards but also helps students follow their passions and become self-directed learners.

Students explore information daily in great depth via projects, field experiences, problem solving, collaborative groups, oral and written communication, entrepreneurialism, and much more, using a skill mastery approach. This approach creates a dynamic learning experience that allows independent thinking and problem solving to develop and flourish.

Our Program

Our curriculum includes:

- Core
  - A core subjects curriculum based on Texas' TEKS standards
- Choice
  - A personalized electives curriculum designed from the input of our students and parents
- Exposure
  - An experiential curriculum highlighting things and events all children should be able to experience without regard to social class or income
- AFIA Curriculum Design

- K–5 Design
- 6–8 Design
- 9–12 Design

In addition to our regular curriculum, students receive unique learning opportunities through our community mentorship program that reflect our four interlocking areas of growth. At Anne Frank, we work hard to:

- Become an expert learner (with problem solving and creative thinking skills)
- Develop leadership skills
- Become a person of principle and character
- Make the world a better place through service

Our Facility

Our schools are designed to create a small, family-like environment that allows each student and teacher to know one another and ensure that no one "falls through the cracks." By partnering with award-winning, international architects at Fielding Nair, we've designed an incredible campus that includes eight outdoor learning areas (including a tree house and amphitheater); a variety of creative, indoor learning studios (like MakerSpaces and Yoga studios); and zero hallways! All of our facilities are also equipped with seamless technology via SkyDrives.

Our Mission & Vision

Mission: To increase the capacity for human greatness.
Vision: Creating 21st century learning models for use around the world through leadership, innovation, safety, technology, integration, synergy, and transformative power.

Our Core Values

Value gives meaning to what we do. But at Anne Frank, our core values of **innovation**, **embracing greatness**, **integrity**, and **joy** are the driving force behind every decision we make.
At Anne Frank, we are active members of our community—**We Belong**. We are becoming experts in every life arena—We can **Be Great**. And we know there is purpose in life—Therefore, we choose to **Find Joy**. In short, we believe students can be self-directed learners, and schools can be both fun and challenging.

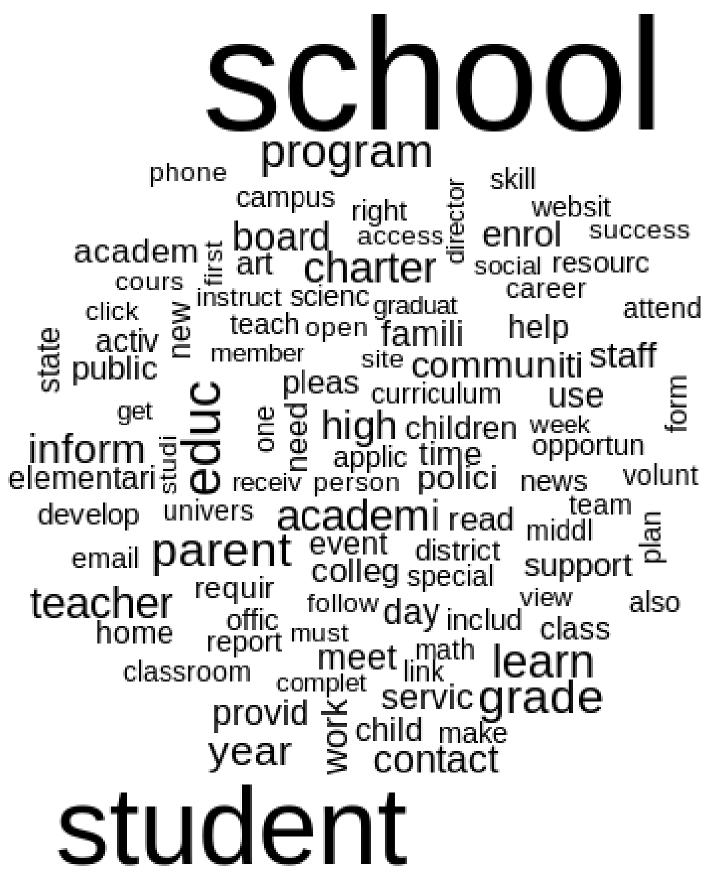


Figure 4.1: *Word cloud of ritual words across corpus*. By virtue of very frequent use, these words are not context-specific; they contrast with ideological terms reflecting community-specific educational beliefs and ideals. *Note:* The text was preprocessed by removing punctuation, stopwords, and numbers; stemming words using the Porter stemmer; and removing infrequent words (those that occur less than 30 times in total).

Table 4.1. Descriptive Statistics and Correlations.

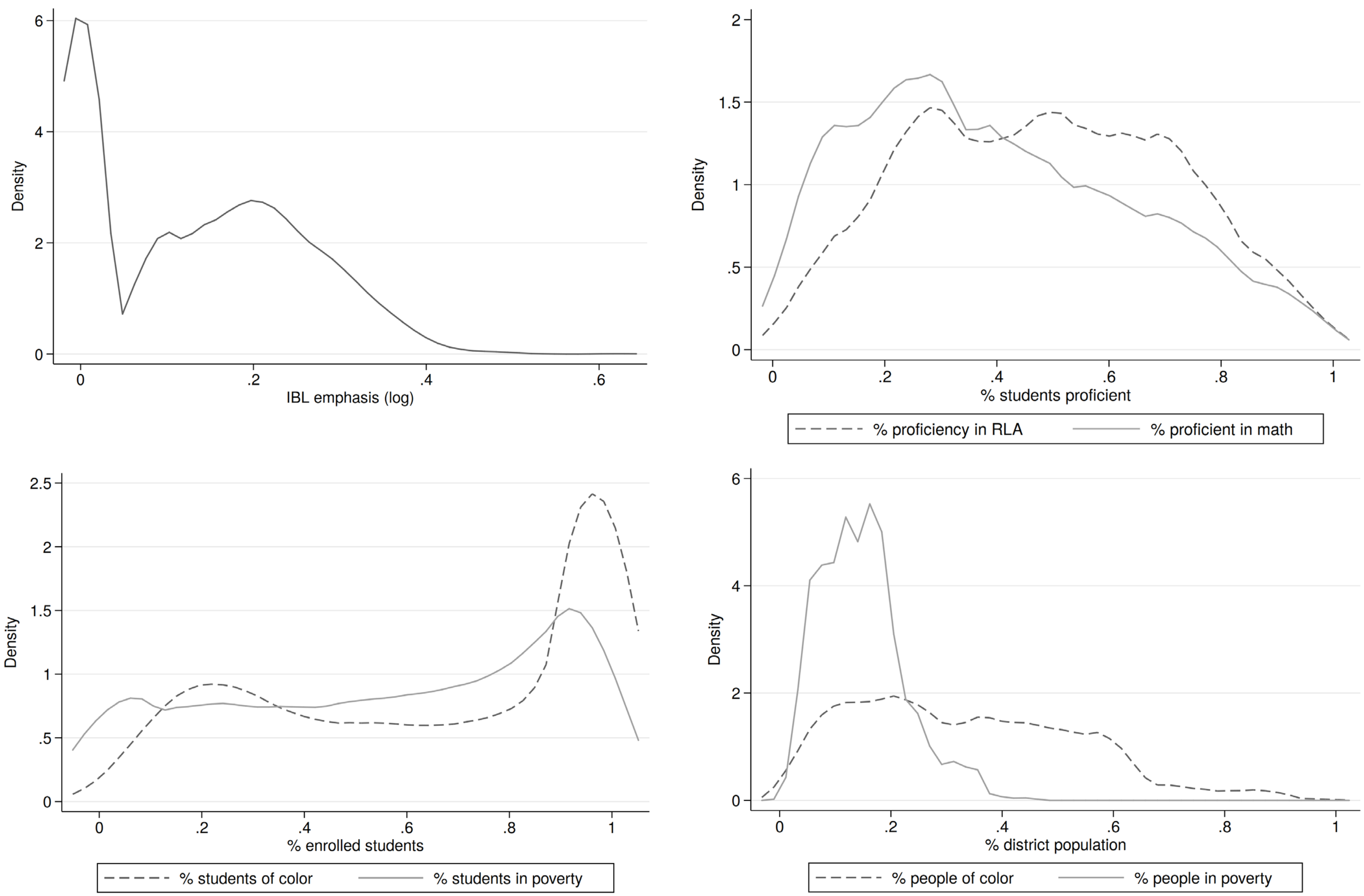
	IBL	RLA	math	pov_ schl.	SOC	pov_ SD	POC _SD	pri- mary	mid- dle	high	other	age (log)	size (log)	urb- an	% PDF	# wor- ds (K)	RLA blur[b]	math blur[b]
Mean	.133	.483	.402	.555	.650	.147	.335	.467	.099	.219	.215	1.73	5.64	.571	.005	23.8	6.97	7.33
Median	.129	.470	.370	.597	.739	.141	.311	.000	.000	.000	.000	1.95	5.75	1.00	.000	8.81	5.00	5.00
Standard error	.120	.231	.245	.325	.324	.075	.200	.499	.299	.413	.411	.969	.968	.495	.047	142	10.4	10.8
IBL emphasis	1																	
% proficiency RLA	.10*	1																
% proficiency math	.05*	.80*	1															
% student poverty	-.18*	-.35*	-.34*	1														
% students of color	-.17*	-.28*	-.17*	.53*	1													
% district poverty	-.11*	-.13*	-.11*	.42*	.47*	1												
% district POC	-.06*	-.18*	-.13*	.27*	.62*	.56*	1											
Primary school [a]	.06*	.02	.11*	-.04*	.05*	.01	.07*	1										
Middle school [a]	-.06*	-.02	.01	.10*	.12*	.02	.07*	-.31*	1									
High school [a]	-.06*	.00	-.09*	.01	.00	.02	-.05*	-.50*	-.18*	1								
Other grade range	.04*	-.01	-.05*	-.04*	-.14*	-.04*	-.08*	-.49*	-.17*	-.28*	1							
Years open (log)	.01	.06*	.01	-.02	-.11*	.00	-.07*	.02	-.10*	.03*	.03*	1						
# students (log)	.09*	.17*	.14*	-.03*	.16*	.06*	.18*	.08*	-.07*	-.20*	.16*	.17*	1					
Urban locale	-.03*	-.07*	-.01	.23*	.46*	.38*	.43*	.03	.08*	.01	-.10*	-.06*	.08*	1				
% PDF web pages	.05*	.01	.00	-.02	-.01	-.02	-.02	-.02	.03*	.00	.01	.02	.01	-.03*	1			
# words	.20*	.03*	.04*	-.04*	-.02	-.03*	.00	-.02	-.01	.01	.02	-.03*	-.01	-.01	.06*	1		
RLA blurring	-.06*	-.09*	-.11*	.00	-.09*	-.03*	-.10*	-.13*	-.10*	.30*	-.07*	-.11*	-.56*	-.05*	-.01	-.02	1	
Math blurring	-.08*	-.08*	-.08*	.01	-.08*	-.02	-.11*	-.15*	-.11*	.34*	-.08*	-.11*	-.56*	-.05*	-.02	-.02	.93*	1

Sources: American Community Survey 2012-16 (U.S. Census Bureau 2018), Common Core of Data 2015-16 (NCES 2019), EdFacts Achievement Results for State Assessments (USDE 2018), and the author's data collection and calculations.

NOTES: [a] Binary indicators of grade range served; the baseline is "other" (incl. ungraded). [b] This indicates the degree of data "blurring" by USDE to protect student groups' privacy. Higher blurring reflects less precise data.

* $p<0.05$

Acronyms: IBL = inquiry-based learning; RLA = reading/language arts; SOC = students of color; POC = people of color; SD = school district; pov = poverty



Figures 4.2 – 4.5. *Univariate distributions for key variables: IBL, academic proficiency, and school/school district poverty and race*. See Table 4.1 for descriptive statistics and correlations.

Table 4.2: Counts and similarities for IBL dictionary terms.
Notes: The 5 **bolded** (seed) terms were used to generate the full 50-term dictionary. The 15 terms in *italics* are part of the 20-term narrow IBL dictionary (together with the seed terms).

Term	Frequency in corpus	Cosine similarity to seed terms
inquiry-based	1258	0.8504
problem-based	198	0.8673
discovery-based	29	0.8212
experiential	1916	0.8053
constructivist	251	0.8219
hands-on	48423	0.6635
problem-solving	7125	0.5836
critical thinking	6647	0.6942
real-world	6397	0.6002
project-based	4129	0.7535
real-life	2209	0.5789
interdisciplinary	1565	0.6910
student-centered	1468	0.6772
critical thinkers	1061	0.6236
expeditions	675	0.5807
child-centered	654	0.7195
experimentation	474	0.6129
student-based	355	0.7364
immersive	295	0.6753
activity-based	260	0.7891
student-driven	238	0.7030
intrinsically	225	0.7452
reality-based	196	0.7220
multi-disciplinary	173	0.7260
learner-centered	125	0.7720
interest-based	118	0.7576
minds-on	117	0.7786
metacognitive	111	0.6791
integrative	109	0.7554
experience-based	102	0.7552
multi-dimensional	99	0.7053
constructivism	97	0.7381
student-directed	74	0.7041
choice-based	58	0.7599
project-oriented	45	0.7552
intellectually stimulating	42	0.7255
inquiry/research	38	0.7239
field-based	37	0.7378
project-focused	36	0.7200
stimulate critical	36	0.7296
student-centric	36	0.7528
active inquiry	26	0.8035
inquiry-driven	25	0.8068
child-directed	24	0.7756
child-initiated	23	0.7571
activity/project	23	0.7075
experientially	20	0.7862
socratic	14	0.7994
problem-centered	12	0.7787
project-centered	10	0.8219

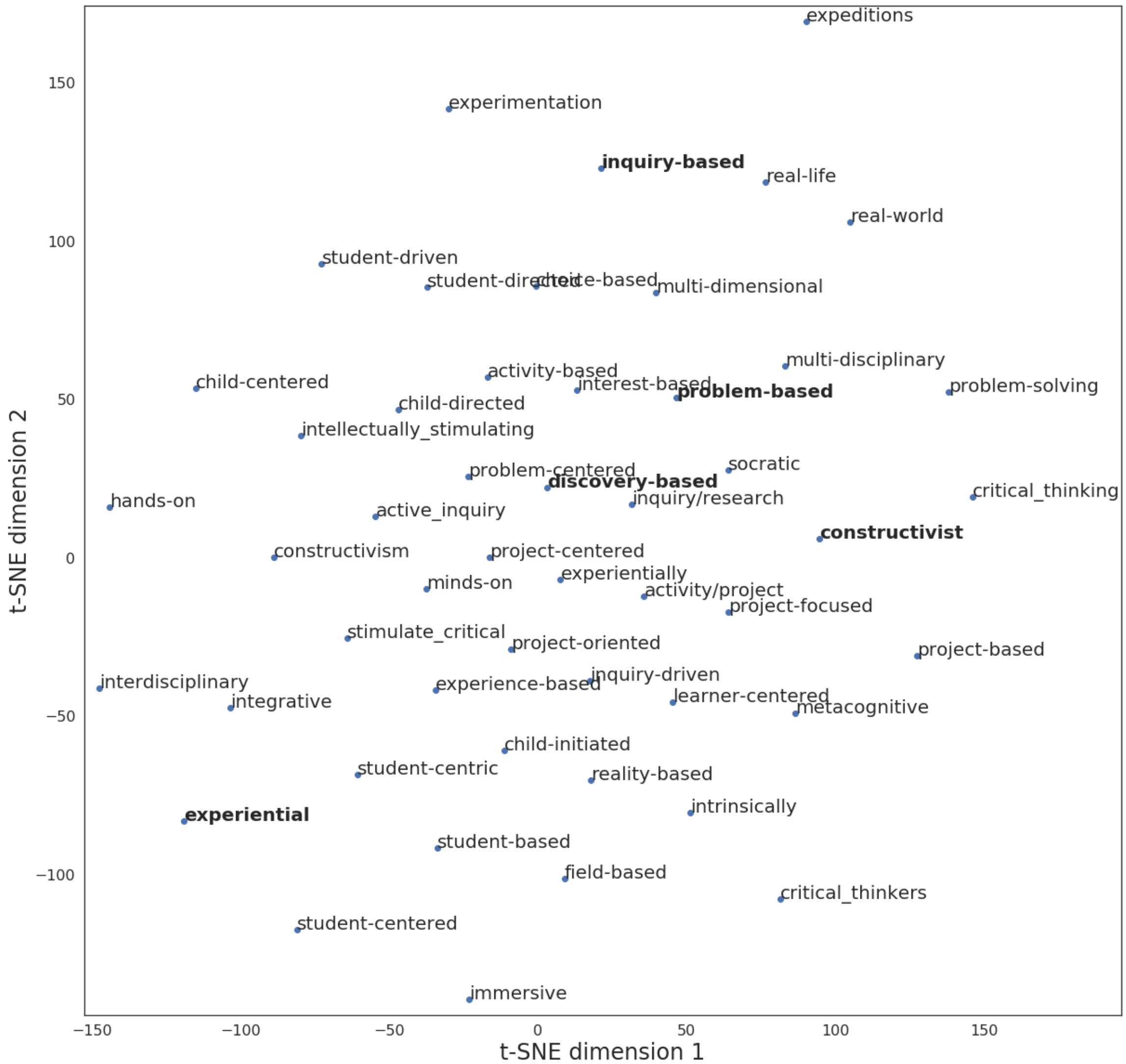


Figure 4.6. *Visualization of IBL dictionary in vector space*. This is a projection of the 300-dimension word embedding vector space into two dimensions through t-Distributed Stochastic Neighbor Embedding (t-SNE), a non-parametric means for visualizing high-dimensional data (Maaten and Hinton 2008). The 50 points represent word vectors from the dictionary for inquiry-based learning (IBL), with the five seed terms in **bold.** Two terms overlap above the graph's center: these are "student-directed" with "choice-based". Pointwise positions are based on cosine distances and preserve local structure, such that points close in high-dimensional space remain close in low-dimensional space. The axes are an artifact of dimensionality reduction and are not directly interpretable. (See text for explanation of cosine similarity scores and dictionary development.)

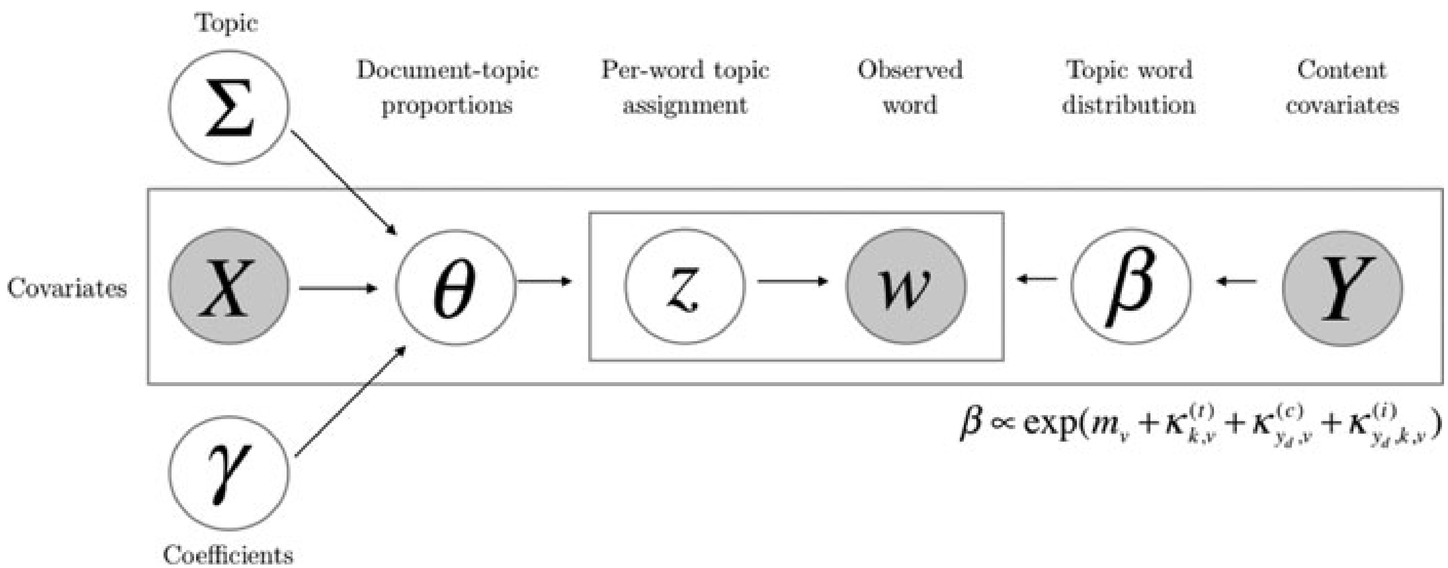


Figure 5.1: Illustration of the Structural Topic Model. *Source*: Roberts, Stewart, and Airoldi 2016:990.

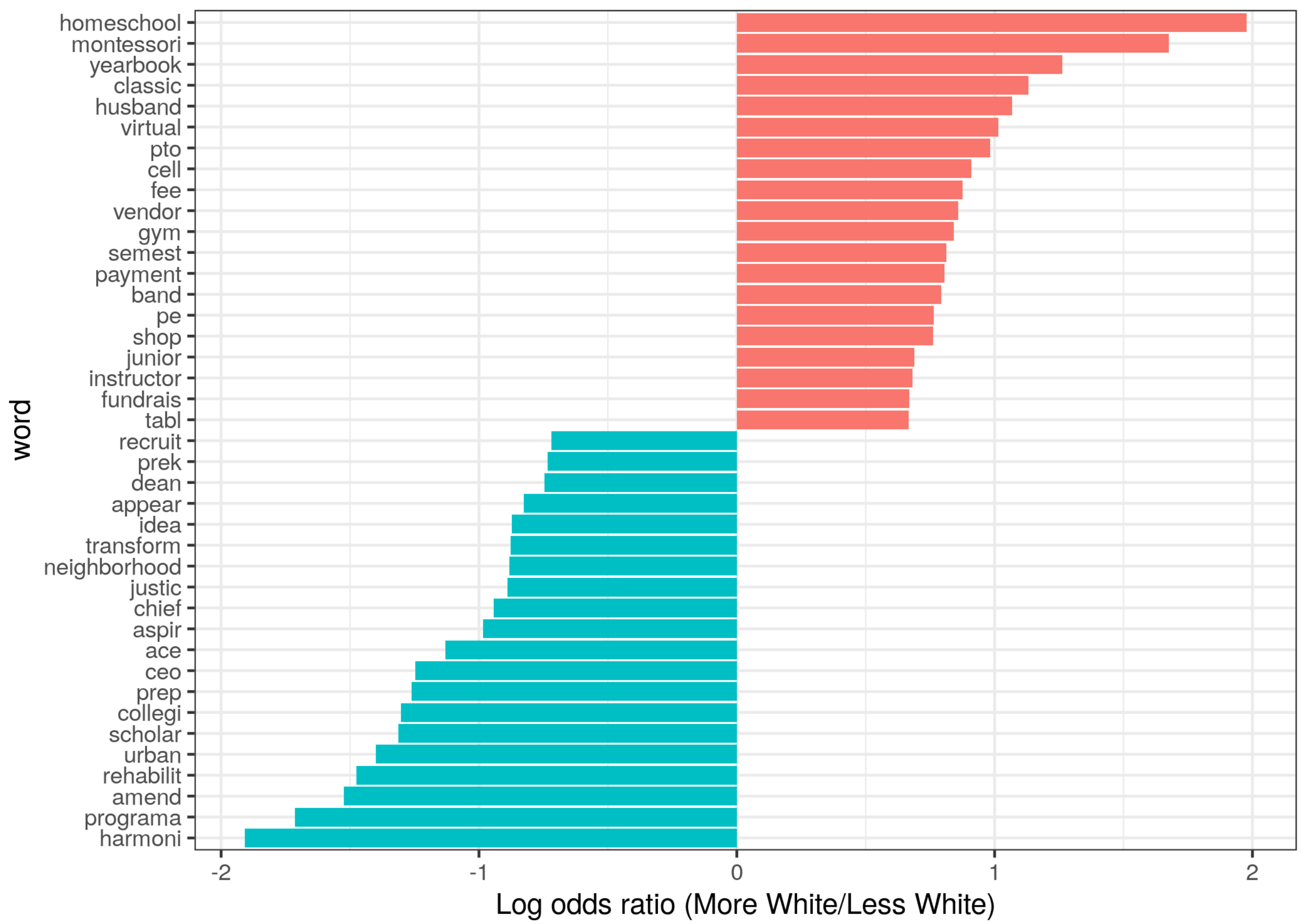


Figure 5.2: *Distinctive words by race.* Schools below median proportion white students are to the left of the center line, while schools above median proportion white students are to the right. *Note*: The text was preprocessed in the same manner as for Structural Topic Models: words were lower-cased and stemmed with the Porter stemmer; punctuation, stopwords, and numbers were removed; and infrequent words (occurring fewer than 30 times across the corpus) were filtered out.

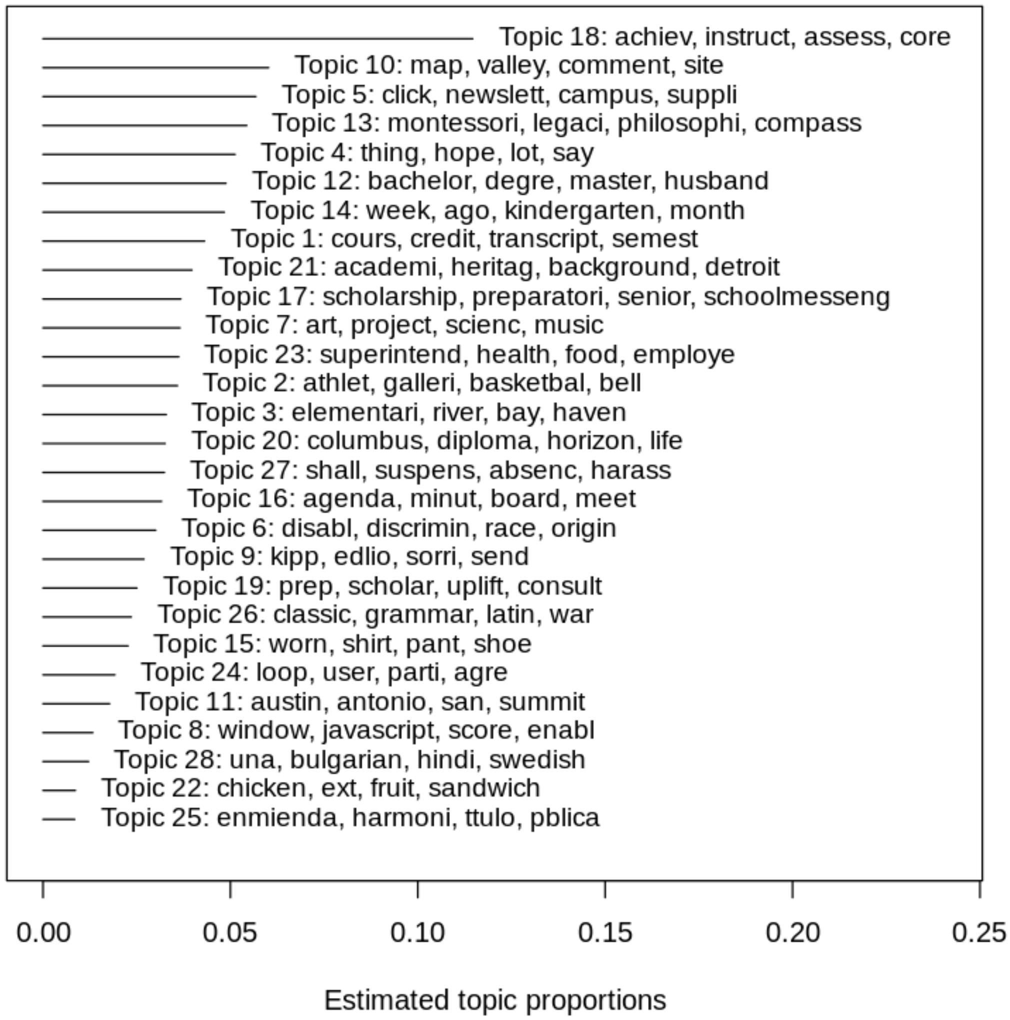


Figure 5.3: *Topic proportions and top words per topic.* Listed next to each topic are the four top words as ranked by a simplified frequency-exclusivity (FREX) metric, which balances the probability of a word occurring in a given topic with (slightly heavier emphasis on) the exclusivity of the word to that topic (Bischof and Airoldi 2012; Roberts et al. 2014; Roberts, Stewart, and Tingley 2019).

Table 5.1. Linear Regressions Predicting Topic-Proportions with School and School District Poverty and Race.

Dependent variable:[a, b] Topic number and label[c]	% pov. schl. (no controls)[d]	% pov. schl. (w. controls)[d]	% SOC (no controls)	% SOC (w. controls)	% pov. SD (no controls)	% pov. SD (w. controls)	% POC SD (no controls)	% POC SD (w. controls)
1: Course Requirements	-0.000334***	-0.000292***	-0.0402***	-0.0345***	-0.00106***	-0.000871***	-0.000575***	-0.000419***
2: Athletics	0.000106**	0.0000805*	0.0216***	0.0171***	0.0000578	-0.000159	0.000191**	0.0000932
3: Schl. Names & Places	0.0000629	0.0000271	0.00406	-0.00281	0.000212	0.0000675	0.000153*	0.0000732
4: Conversation	-0.000164***	-0.000171***	-0.0215***	-0.0229***	-0.000835***	-0.000875***	-0.000193***	-0.000154*
5: Web Administration	-0.000102*	-0.0000746	-0.00687	-0.00806	-0.000477**	-0.000395*	-0.000125	-0.000164*
6: Civil Rights	0.0000055	0.0000124	0.00655	0.00494	0.000354	0.000385	0.000122	0.0000688
7: Inquiry-Based Lrng.	-0.000191***	-0.000207***	-0.0212***	-0.0282***	-0.000476***	-0.000599***	-0.0000592	-0.0000761
8: Web Interface	-0.0000681**	-0.0000451	-0.00776**	-0.00471	-0.000159	-0.0000362	-0.0000592	0.0000100
9: Online Access	0.000363***	0.000308***	0.0449***	0.0385***	0.000573*	0.000215	0.000543***	0.000455***
10: Communication	0.0000177	0.0000861	-0.00979*	0.00148	-0.000241	0.0000716	-0.000251***	-0.000104
11: Urban Locales	0.000369***	0.000361***	0.0401***	0.0397***	0.00165***	0.00175***	-0.000394***	-0.000647***
12: Faculty and Staff	-0.0000548	-0.0000821	-0.00196	-0.00930	-0.000548**	-0.000749***	0.0000760	0.00000784
13: Holistic Education	-0.000373***	-0.000352***	-0.0462***	-0.0400***	-0.00145***	-0.00132***	-0.000308***	-0.000133
14: Schedules/Calendars	-0.000326***	-0.000293***	-0.0296***	-0.0289***	-0.000862***	-0.000680***	-0.000236***	-0.000207**
15: Dress Code	-0.000253***	-0.000207***	-0.00109	0.00429	-0.000395*	-0.000138	-0.0000821	-0.0000304
16: Schl. Governance	-0.0000938**	-0.0000806*	-0.0101**	-0.00415	-0.000280	-0.000136	-0.000133*	-0.0000366
17: College Prep	0.000148***	0.000101**	0.0195***	0.0153***	0.000498**	0.000231	0.000254***	0.000180**
18: Standards/assesmnt.	0.000337***	0.000325***	0.0282***	0.0346***	0.000196	0.000154	0.000366***	0.000428***
19: Virtual Education	0.000237***	0.000169***	0.0253***	0.0117*	0.000216	-0.000183	0.000341***	0.000115
20: Diploma/Graduation	0.000183***	0.000111*	0.0171***	0.0129*	0.00182***	0.00166***	0.0000984	0.0000855
21: Parent Involvement	0.000221***	0.000273***	0.0130**	0.0131*	0.00134***	0.00148***	0.000427***	0.000439***
22: Meals	-0.0000556*	-0.0000453	-0.00477*	-0.00324	-0.000263**	-0.000205*	-0.0000160	0.0000113
23: Services	0.000159***	0.000203***	-0.00305	0.0115**	0.000394*	0.000634***	-0.0000615	0.000132*
24: Terms of Agreement	-0.000122*	-0.000115*	-0.0180***	-0.0196***	-0.0000113	0.0000657	0.0000202	0.0000729
25: Spanish Language	0.0000842*	0.0000426	0.0171***	0.0111**	0.000188	-0.0000123	0.00000987	-0.000148*
26: Classical/Lib. Arts	-0.00028***	-0.000291***	-0.0258***	-0.0268***	-0.000726***	-0.000755***	-0.000160*	-0.000128
27: Student Discipline	-0.000104*	-0.0000676	-0.0142**	-0.00823	-0.0000564	0.000214	-0.000167*	-0.0000761
28: International	0.000229***	0.000223***	0.0247***	0.0252***	0.000354	0.000199	0.000217**	0.000146

Sources: American Community Survey 2012-16 (U.S. Census Bureau 2018), Common Core of Data 2015-16 (NCES 2019), and the author's data collection and calculations. *Acronyms*: SOC = students of color; POC = people of color; SD = school district; pov = poverty

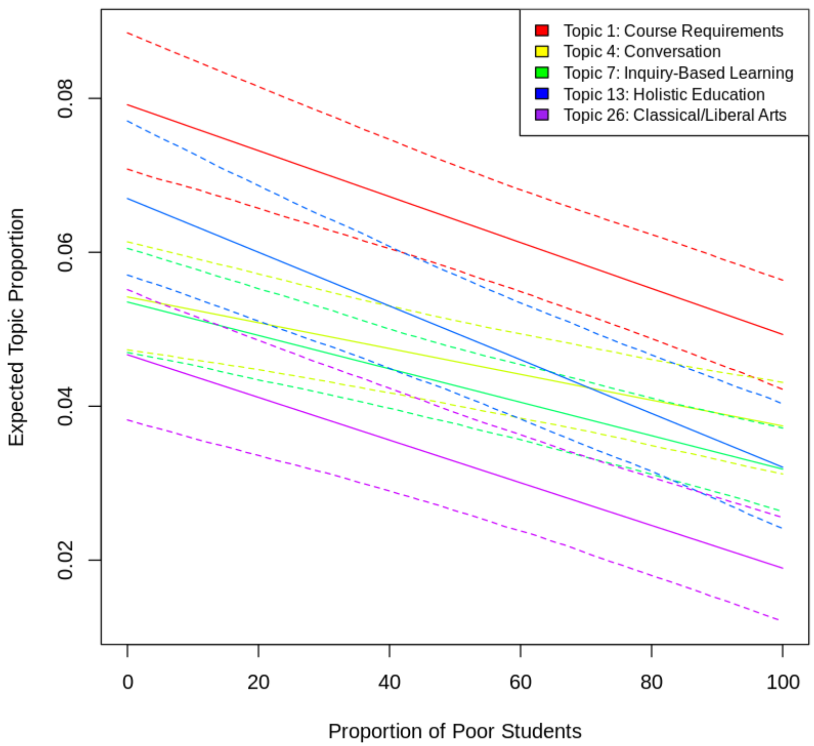
Notes: [a]Topic labels chosen by me. [b]For readability, I omit controls and standard errors and list the outcome variable in rows and predictors in columns. [c]The outcome is the proportion for a given topic in a given document. Each topic proportion is calculated in a separate linear model for each predictor (yielding 8 models per topic); thus, each cell represents a beta coefficient from its own model. [d]Controls included school level (primary, middle, high; each is binary), years open (logged), number students (logged), urban locale (binary), and % PDF web pages.

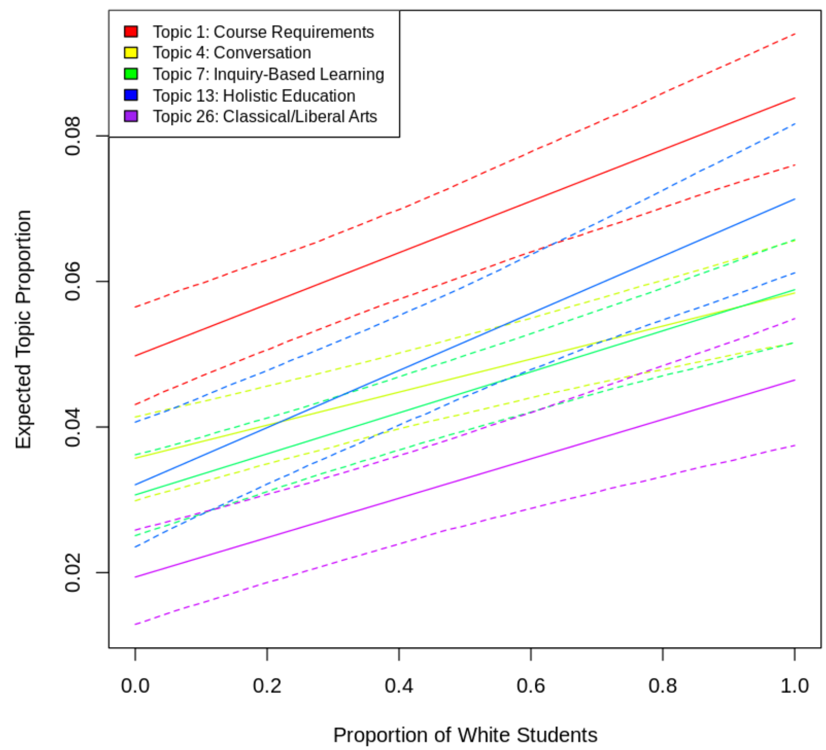
*** p<0.001, ** p<0.01, * p<0.05

Table 5.2. Clusters of Topics with their 20 Most Distinctive Words and Associations with Race and Class.

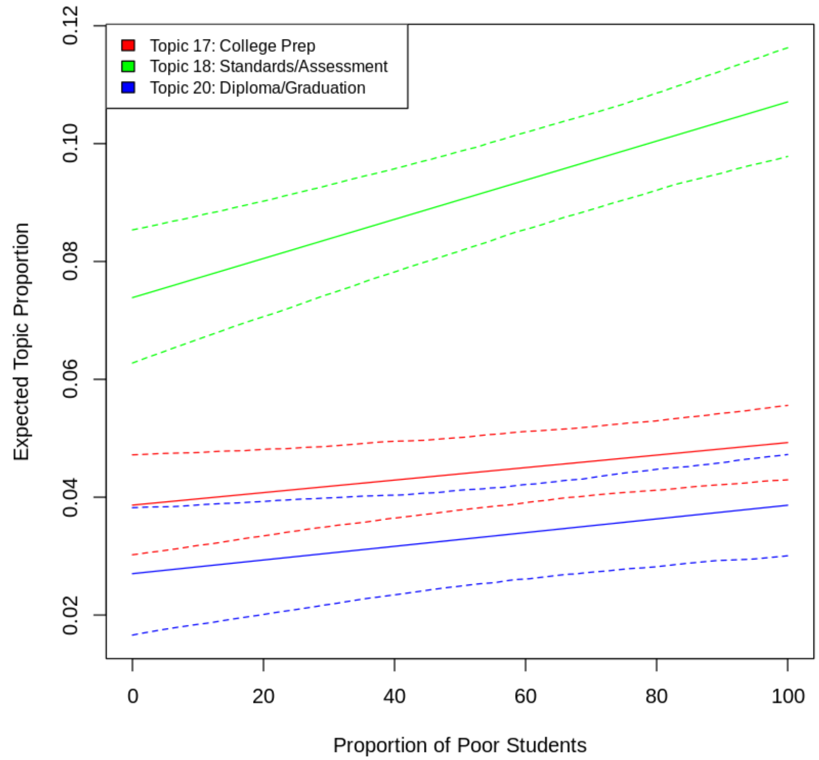
Cluster	*Academics*[a]					*Standards/College-Bound*			*Serving the Disadvantaged*			
Race/ Class Associations[b]	Strongly (p<0.001), robustly associated with more *white* or *affluent* schools and districts. Only the association with district ethnic composition is sensitive to controls (for some topics).					Strongly (p<0.001) associated with *students of color* or school *poverty*. Mixed, positive district effects; some robustness.			Strongly (p<0.001) associated with *students of color* or school *poverty*. Mixed but positive district effects; some robustness to controls.			
Topic	1: Course Requirements[c]	4: Conversation	7: Inquiry-Based Learng	13: Holistic Education	26: Classicl/ Liberal Arts	17: College Prep	18: Standards / Assessment	20: Diploma / Graduation	2: Athletics	11: Urban Locales	19: Virtual Education	23: Services
Most Distinctive Words	cours[d, e]	thing	art	montessori	classic	scholarship	achiev	diploma	athlet	austin	prep	superintend
	credit	hope	project	legaci	grammar	preparatori	instruct	life	galleri	antonio	scholar	health
	transcript	lot	scienc	philosophi	latin	senior	assess	columbus	basketbal	san	uplift	food
	semest	say	music	compass	war	presenc	core	skill	bell	summit	consult	employe
	test	want	artist	children	write	corpor	improv	akron	club	idea	virtual	youth
	onlin	realli	stem	garden	geographi	award	valu	high	ace	gateway	online	nutrit
	algebra	even	visual	donat	vocabulari	colleg	goal	career	boy	vista	forgot	district
	elect	much	creativ	farm	literatur	cadet	develop	locat	girl	diego	memphi	salari
	requir	alway	engin	inspir	word	alumni	standard	enrol	basebal	region	jersey	fund
	colleg	said	robot	waldorf	topic	collegi	progress	graduat	volleybal	frontier	cell	evalu
	english	know	theater	approach	text	internship	expect	cleveland	denver	jose	urban	agenc
	graduat	like	danc	preschool	analyz	dalla	level	delta	soccer	rio	oop	fiscal
	complet	everyon	perform	natur	understand	gpa	rigor	flexibl	photo	pike	face--fac	expenditur
	offer	great	math	gift	sequenc	financi	ensur	earn	footbal	headquart	chat	mental
	technic	never	design	classroom	earth	aid	strong	excel	advisori	pedro	info	insur
	exam	love	theatr	environ	sentenc	graduat	knowledg	dayton	varsiti	blvd	newark	facil
	sat	better	explor	beauti	ancient	honor	model	premier	sport	leader	someth	treatment
	advanc	feel	connect	sustain	equat	mentor	growth	testimoni	transpar	vocat	enrol	tax
	class	favorit	studi	emot	knowledg	campus	strategi	path	handbook	francisco	talk	medic
	placement	friend	technolog	foster	poetri	foundat	engag	futur	schedul	transit	blend	servic

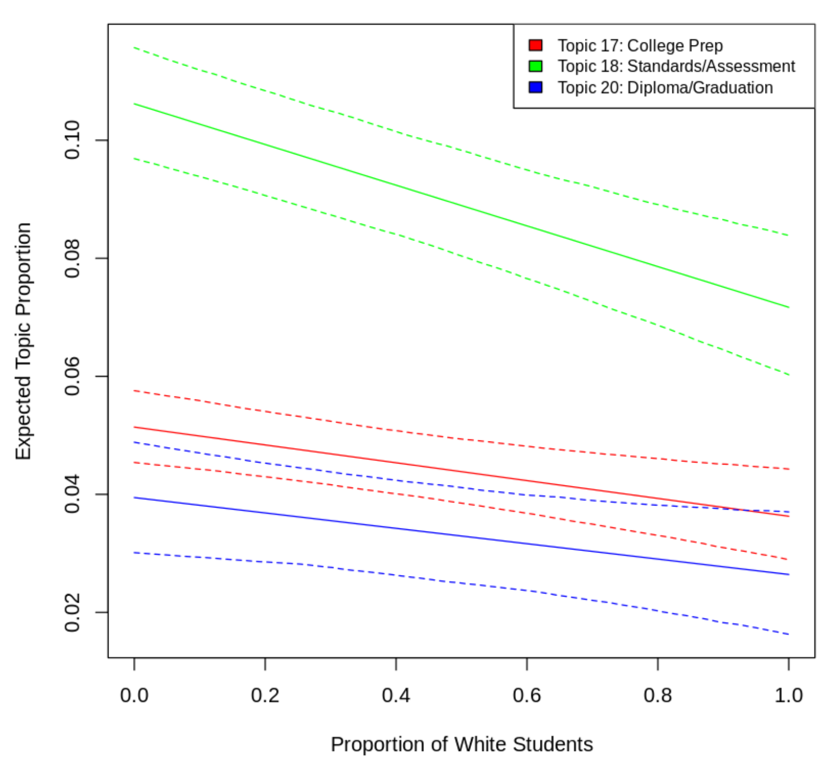
Notes: [a]I inductively derived these (selected) clusters by examining patterns in topics' top words and associations with race and class. [b]See above table. [c]I chose these topic labels by examining top words per topic and representative documents. [d]The words are those most distinctive to each topic, as measured by a simplified frequency-exclusivity (FREX) metric (Roberts et al. 2014). [e]Words were stemmed using the Porter stemmer.



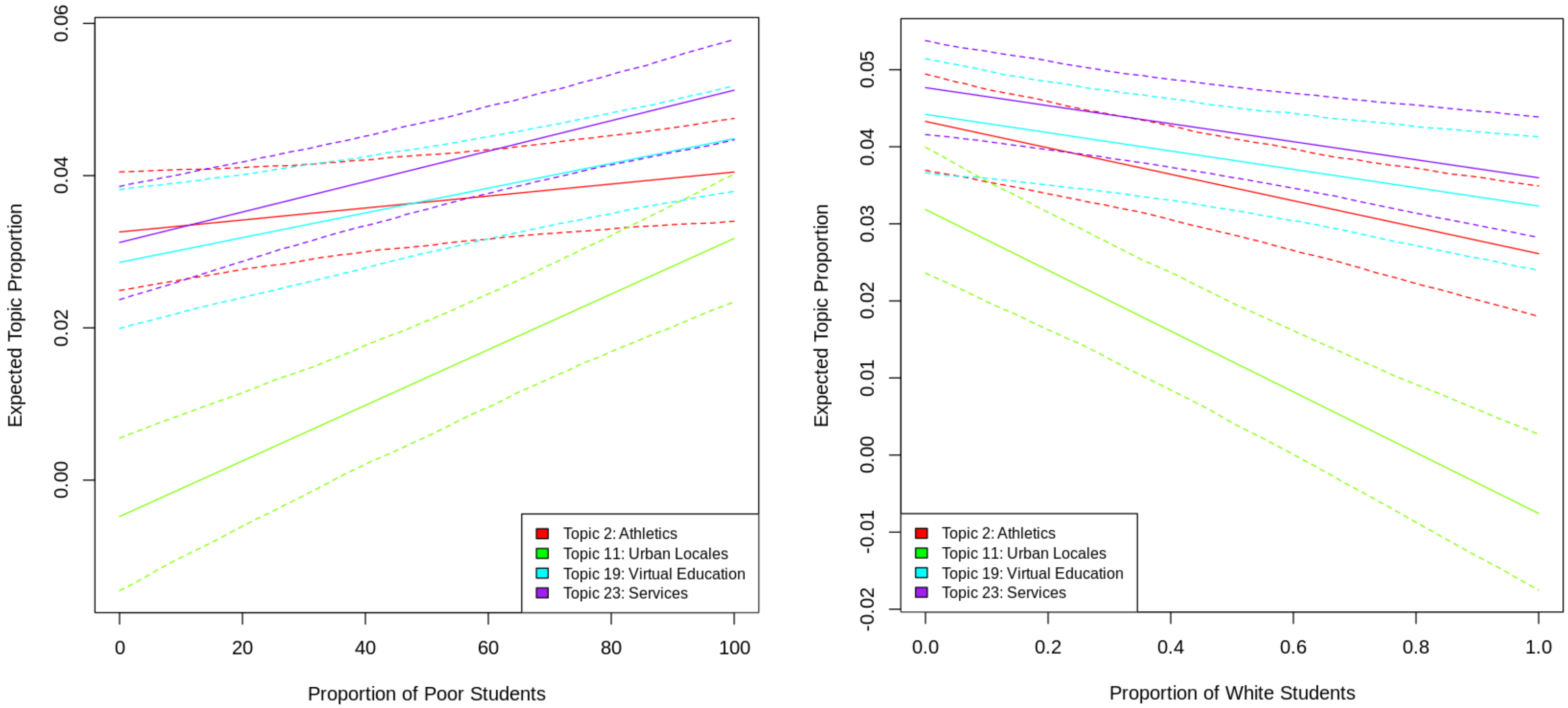


Figures 5.4 and 5.5: *Results of regressing proportions of topics in* Academics *cluster on school proportion in poverty or of white ethnicity*. See also Tables 4.1 and 4.2.

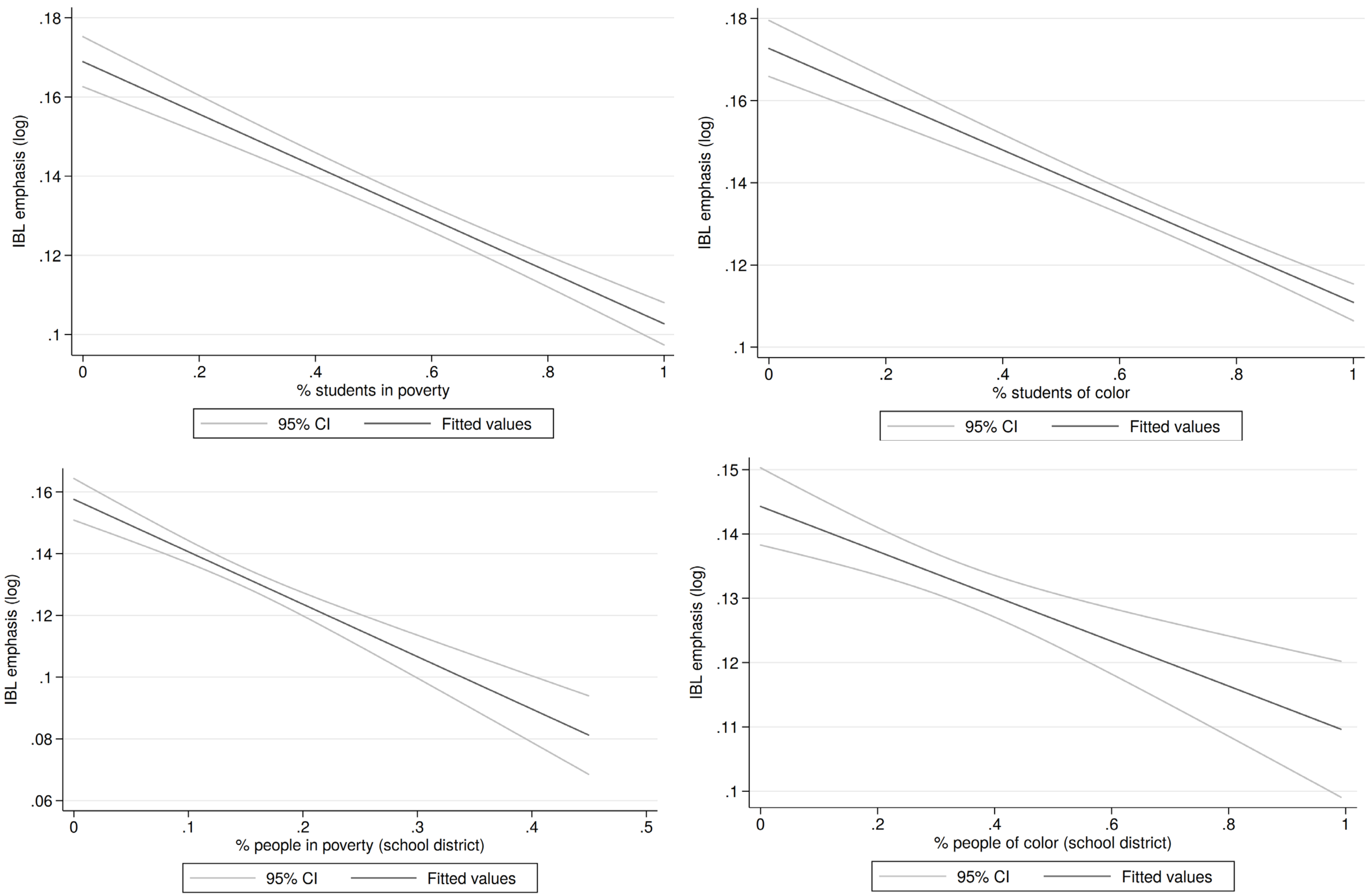




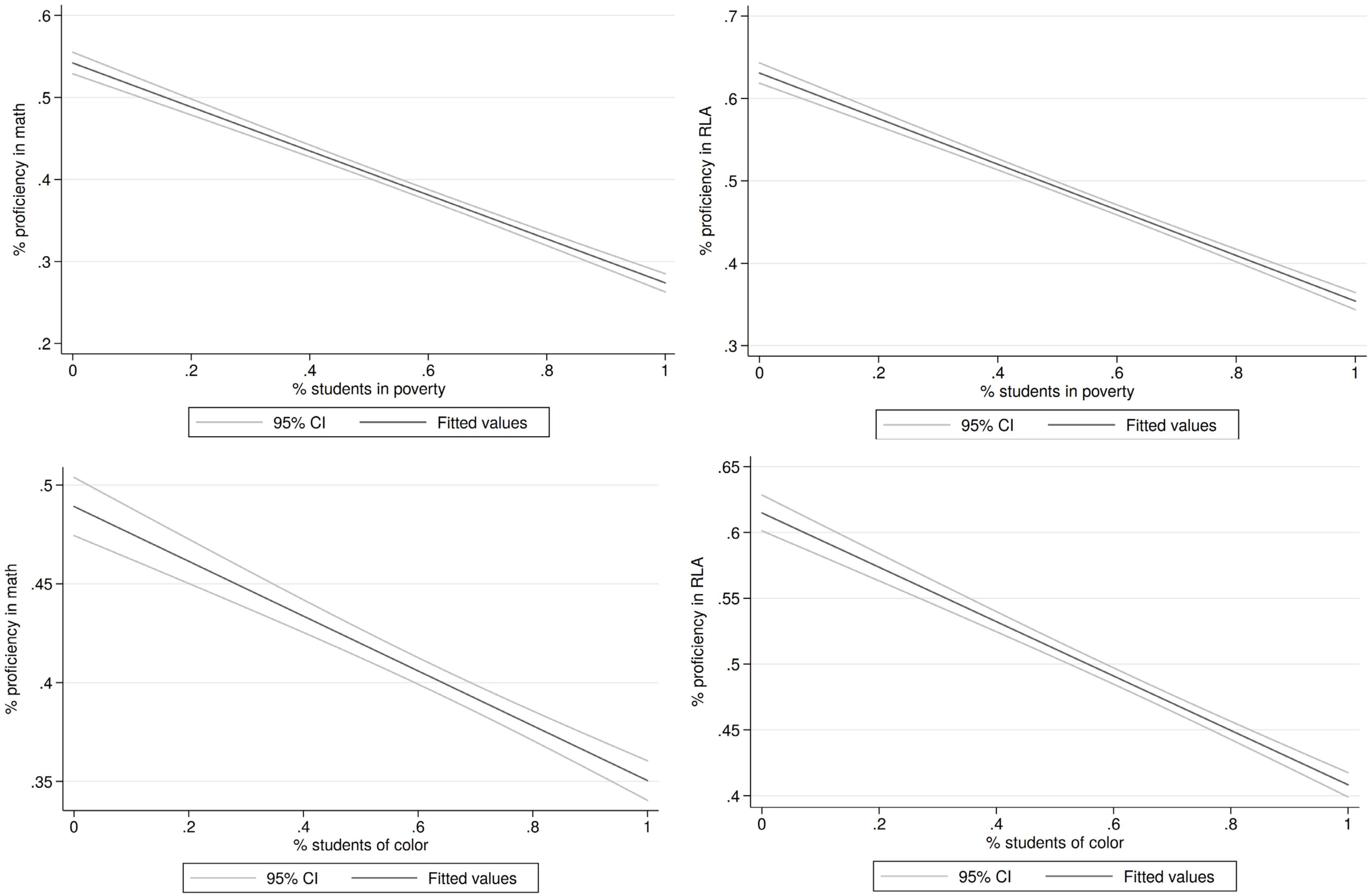
Figures 5.6 and 5.7: *Results of regressing proportions of topics in* Standards/College-Bound *cluster on school proportion in poverty or of white ethnicity*. See also Tables 4.1 and 4.2.



Figures 5.8 and 5.9: *Results of regressing proportions of topics in* Serving the Disadvantaged *cluster on school proportion in poverty or of white ethnicity*. See also Tables 4.1 and 4.2.



Figures 6.1 – 6.4. *Lines of best fit between IBL emphasis (outcome) and school/school district poverty and race (predictors)*. Lacking controls and nested relationships, these simple models correspond to the correlations in Table 4.1.



Figures 6.5 – 6.8. *Lines of best fit between proficiency in math and reading/language arts (outcome) and school poverty and race (predictors)*. See Table 4.1 for corresponding correlations. Similar lines are observed for school district poverty and race.

Table 6.1. Mixed Linear Regressions Predicting IBL Emphasis with School and School District Poverty and Race.

Outcome: IBL emphasis Independent variable	Model 1a: Controls only	Model 1b: School poverty	Model 1c: School race	Model 1d: School district poverty	Model 1e: School district race
Poverty and race					
% students in poverty		-0.000617***			
		(5.00e-05)			
% students of color			-0.0733***		
			(0.00539)		
% district in poverty				-0.212***	
				(0.0221)	
% district people of color					-0.0302***
					(0.00869)
School controls					
Primary school[a]	0.000856	0.000188	0.00602	0.000415	0.00162
	(0.00394)	(0.00389)	(0.00390)	(0.00391)	(0.00395)
Middle school[a]	-0.0177**	-0.0155**	-0.00918	-0.0183**	-0.0164**
	(0.00590)	(0.00582)	(0.00584)	(0.00585)	(0.00590)
High school[a]	-0.0134**	-0.0137**	-0.00784	-0.0125**	-0.0129**
	(0.00472)	(0.00466)	(0.00467)	(0.00469)	(0.00472)
Years open (log)	-0.00396*	-0.00370*	-0.00607***	-0.00376*	-0.00441**
	(0.00163)	(0.00161)	(0.00161)	(0.00162)	(0.00163)
Number students (log)	0.00916***	0.00801***	0.0124***	0.00953***	0.0102***
	(0.00170)	(0.00168)	(0.00169)	(0.00169)	(0.00172)
Urban locale (binary)	0.000541	0.00850**	0.0188***	0.0117***	0.00503
	(0.00309)	(0.00312)	(0.00333)	(0.00328)	(0.00335)
% PDF web pages	0.115***	0.112***	0.114***	0.113***	0.114***
	(0.0319)	(0.0315)	(0.0314)	(0.0316)	(0.0319)
Model parameters					
Constant	0.0661***	0.106***	0.0908***	0.0905***	0.0692***
	(0.0116)	(0.0120)	(0.0116)	(0.0118)	(0.0116)
Variance (σ^2) between CMOs	0.00583	0.00592	0.00570	0.00592	0.00581
Residual variance (σ^2)	0.0117	0.0113	0.0113	0.0115	0.0117
CMO ICC[b] (ρ)	0.343	0.335	0.337	0.340	0.333
Number of CMOs	377	377	377	377	377
Number of observations	5,784	5,784	5,784	5,784	5,784

Model fit					
Log likelihood	4456	4536	4547	4502	4462
Degrees of freedom	10	11	11	11	11
AIC	-8892	-9050	-9072	-8982	-8902
BIC	-8825	-8976	-8999	-8909	-8829
Wald χ^2 [c]	78.1***	233***	265***	173***	90.3***
RE test: χ^2 [d]	666***	663***	643***	696***	654***

Sources: American Community Survey 2012-16 (U.S. Census Bureau 2018), Common Core of Data 2015-16 (NCES 2019), and the author's data collection and calculations.
Notes: [a] Binary indicators of grade range served; the baseline is "other" (incl. ungraded). [b] The Intraclass Correlation Coefficient (ICC) or *rho* (ρ) here measures nesting of the outcome in charter management organizations (CMOs). [c] The Wald χ^2 statistic tests the null hypothesis that all regression coefficients are zero. [d] This conservative χ^2 test assesses the null hypothesis that all random effects (RE) in the model are zero; thus, significance supports the model. Standard errors in parentheses.
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table 6.2. Mixed Linear Regressions Predicting School Poverty and Race with IBL Emphasis and Academic Achievement.

	Number of students in poverty				Number of students of color			
Independent variable	M2a: Controls only	M2b: IBL emphasis	M2c: Academic achievement	M2d: Fully specified	M3a: Controls only	M3b: IBL emphasis	M3c: Academic achievement	M3d: Fully specified
Ideology and academic quality								
IBL emphasis		-0.298***		-0.204***		-0.292***		-0.217***
		(0.0306)		(0.0286)		(0.0233)		(0.0221)
RLA proficiency[a]			-0.450***	-0.430***			-0.335***	-0.314***
			(0.0302)	(0.0304)			(0.0229)	(0.0229)
Math proficiency			-0.0587*	-0.0678*			-0.0561*	-0.0639**
			(0.0298)	(0.0297)			(0.0232)	(0.0230)
School controls								
Primary school[b]	-0.00429	-0.00318	0.00303	0.00391	0.0448***	0.0455***	0.0518***	0.0522***
	(0.00953)	(0.00946)	(0.00901)	(0.00897)	(0.00724)	(0.00714)	(0.00682)	(0.00675)
Middle school[b]	0.0336*	0.0297*	0.0396**	0.0368**	0.0704***	0.0663***	0.0783***	0.0747***
	(0.0141)	(0.0140)	(0.0134)	(0.0133)	(0.0107)	(0.0105)	(0.0101)	(0.0100)
High school[b]	-0.0110	-0.0135	0.00416	0.00231	0.0566***	0.0541***	0.0659***	0.0638***
	(0.0113)	(0.0112)	(0.0108)	(0.0108)	(0.00860)	(0.00848)	(0.00824)	(0.00816)
Years open (log)	0.00167	0.000751	0.00871*	0.00780*	-0.0159***	-0.0167***	-0.00907**	-0.00992***
	(0.00399)	(0.00397)	(0.00376)	(0.00375)	(0.00306)	(0.00302)	(0.00289)	(0.00286)
Number students (log)	-0.0165***	-0.0131**	-0.000526	0.00120	0.00481	0.00805*	0.0222***	0.0237***
	(0.00446)	(0.00443)	(0.00512)	(0.00511)	(0.00337)	(0.00333)	(0.00390)	(0.00386)
Urban locale	0.0687***	0.0710***	0.0627***	0.0646***	0.107***	0.110***	0.0990***	0.101***
	(0.0110)	(0.0109)	(0.0102)	(0.0101)	(0.00919)	(0.00908)	(0.00859)	(0.00852)
% PDF web pages		0.0590		0.0469		0.104		0.101
		(0.0777)		(0.0726)		(0.0601)		(0.0565)
RLA blurring[c]			-3.86e-05	0.000124			0.000832	0.000962
			(0.000941)	(0.000938)			(0.000768)	(0.000757)
Math blurring[c]			-0.000430	-0.000633			-0.000530	-0.000706
			(0.000929)	(0.000927)			(0.000740)	(0.000730)
Model parameters								
Constant	0.568***	0.591***	0.716***	0.730***	0.432***	0.457***	0.507***	0.523***
	(0.0266)	(0.0265)	(0.0327)	(0.0327)	(0.0353)	(0.0349)	(0.0374)	(0.0371)

Variance (σ^2) between states					0.0304	0.0296	0.0292	0.0287
Variance (σ^2) between school districts	0.0364	0.0347	0.0305	0.0294	0.0400	0.0397	0.0356	0.0356
Residual variance (σ^2)	0.0571	0.0563	0.0482	0.0480	0.0332	0.0322	0.0284	0.0278
State ICC[d] (ρ)					0.294	0.292	0.317	0.315
School district ICC[d] (ρ)	0.382	0.373	0.380	0.372	0.386	0.391	0.345	0.345
Number of states					43	43	43	43
Number of school districts	1,481	1,481	1,481	1,481	1,481	1,481	1,481	1,481
Number of observations	5,784	5,784	5,784	5,784	5,784	5,784	5,784	5,784
Model fit								
Log likelihood	-594	-543	-114	-87.5	638	716	1076	1123
Degrees of freedom	9	11	13	15	10	12	14	16
AIC	1206	1108	255	205	-1255	-1408	-2124	-2215
BIC	1266	1181	341	305	-1189	-1328	-2030	-2108
Wald χ^2 [e]	67.2***	164***	1157***	1219***	234***	397***	1225***	1344***
RE test: χ^2 [f]	1892***	1832***	2038***	1956***	2992***	2944***	3313***	3250***

Sources: American Community Survey 2012-16 (U.S. Census Bureau 2018), Common Core of Data 2015-16 (NCES 2019), EdFacts Achievement Results for State Assessments (USDE 2018), and the author's data collection and calculations.

Notes: [a] School proficiency rates in state assessments of reading/language arts (RLA). [b] Binary indicators of grade range served; the baseline is "other" (incl. ungraded). [c] This indicates the degree of data "blurring" by USDE to protect student groups' privacy. This score ranges from 1 (percentiles reported, e.g. 95% academic proficiency for a school) to 50 (medians reported, e.g. 50-100% proficiency), with higher blurring reflecting less precise data. [d] The Intraclass Correlation Coefficient (ICC) or *rho* (ρ) here measures nesting of the outcome in states and school districts. [e] The Wald χ^2 statistic tests the null hypothesis that all regression coefficients are zero. [f] This conservative χ^2 test assesses the null hypothesis that all random effects (RE) in the model are zero; thus, significance supports the model. Standard errors in parentheses.

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

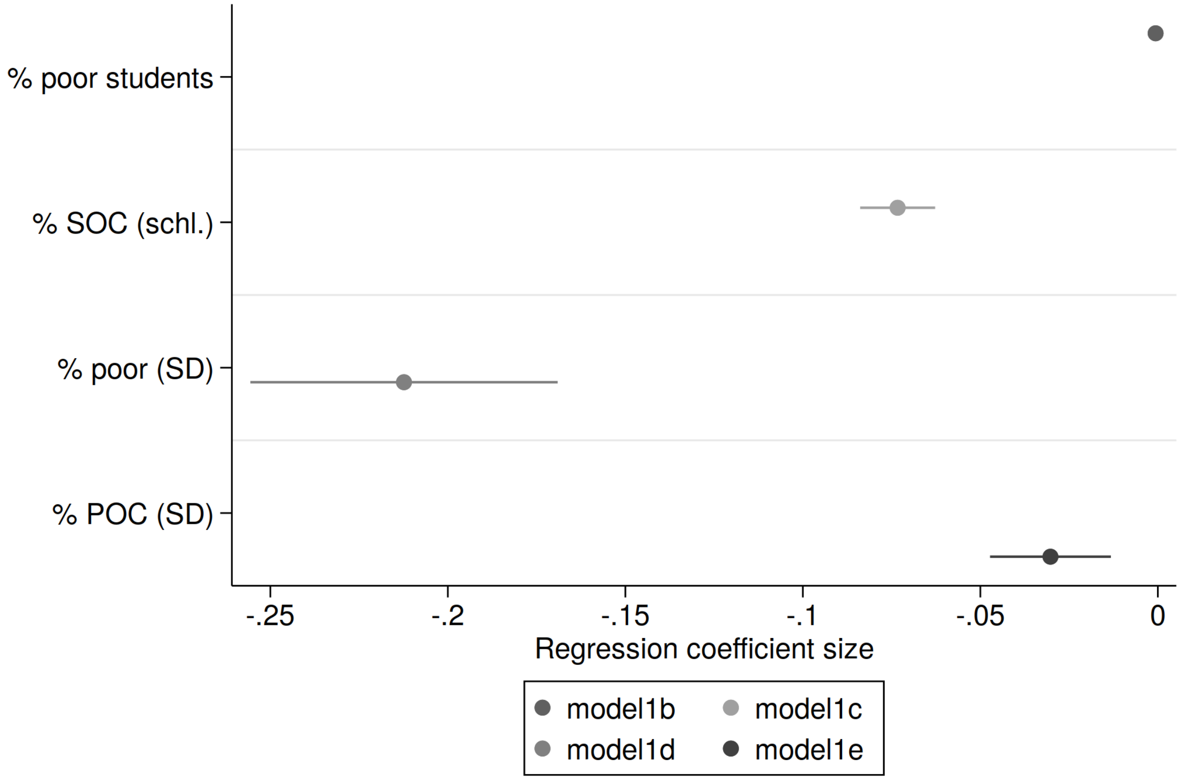


Figure 6.9. *Results of regressing IBL emphasis on school/school district poverty and race (corresponding to table 6.1)*. Each independent variable is modeled separately, and all effects are statistically significant. For clarity, not shown are controls: School level (dummies for primary, middle, high), years open (logged), number of students (logged), urban locale, and % PDF web pages.
Acronyms: IBL = inquiry-based learning; SOC = students of color; POC = people of color; SD = school district.

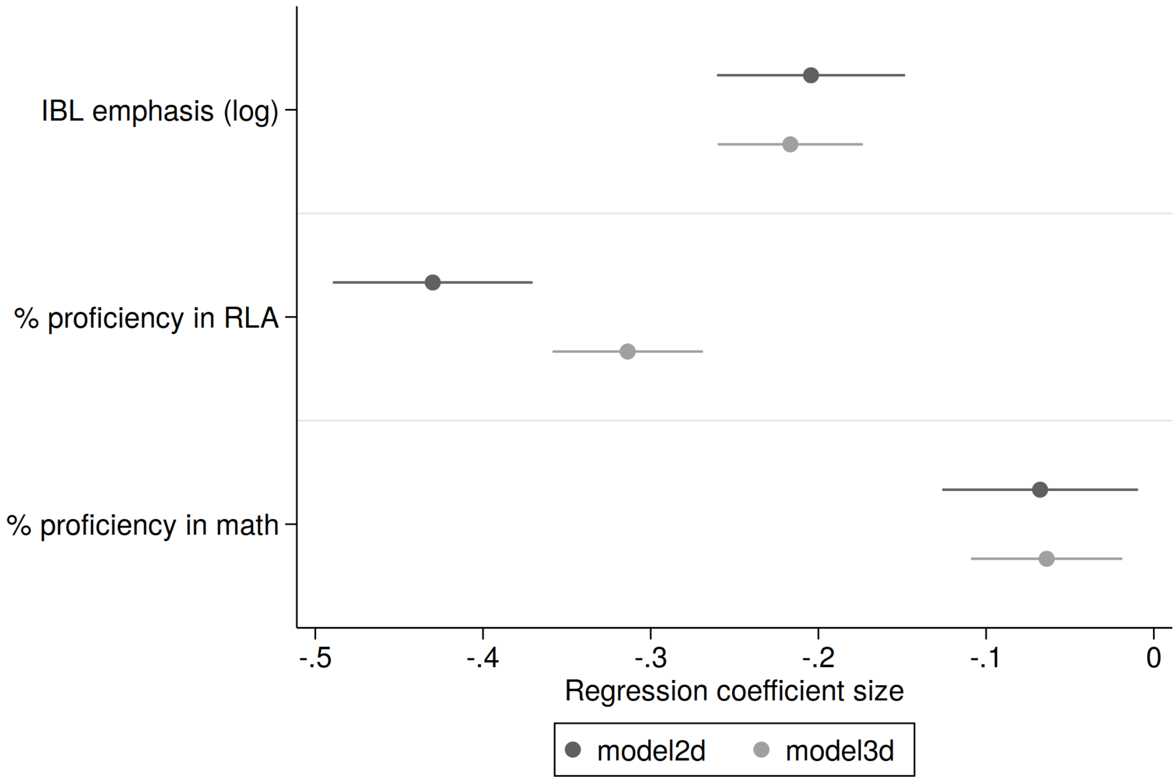


Figure 6.10. *Results of regressing school/school district poverty and race on IBL emphasis and academic proficiency (corresponding to table 6.2)*. The dependent variable of model 2d (upper dot) is % poor students; for model 3d (lower dot), it is % students of color. All effects are statistically significant. For clarity, not shown are controls: School level (dummies for primary, middle, high), years open (logged), number of students (logged), urban locale, percent PDF web pages, and blurring of RLA and math proficiency rates.
Acronyms: IBL = inquiry-based learning; RLA = reading/language arts