The Effects of High Stakes Educational Testing on Enrollments in an Era of Hyper-Expansion: Cross-National Evidence, 1960–2010

Jared Furuta, *Stanford University*
Evan Schofer, *University of California, Irvine*
Shawn Wick, *Central College*

How do national high-stakes exams affect educational expansion across the world? High-stakes exams are conventionally viewed as systems of exclusion that constrain enrollments. In this paper, we situate exams within a broader historical and institutional context and argue that the constraining effect of exams on educational enrollments is a recent phenomenon. Exam systems diffused globally at a time when schooling was a limited enterprise, linked to just a few occupational roles. The later emergence of more inclusive visions of education, culminating in the Education for All (EFA) movement, propelled rapid global educational expansion. In this context, national high-stakes exam systems institutionalize earlier logics of selective education and consequently blunt the impact of more recent expansionary norms. Using panel regression models and a newly constructed dataset of 142 countries from 1960 to 2010, we show that high-stakes exams are associated with lower enrollments. However, this association is strongest in recent years, and exams interact negatively with measures of international pro-educational norms and pressures on nation-states. These findings are consistent with our historical/institutional argument: Exams constrain enrollments in recent years, in part by rendering nations less responsive to global expansionary pressures.

Introduction

How do national high-stakes exams affect educational expansion across the world? Sociologists of education have long recognized the importance of exams as central features of national educational systems (e.g., Allmendinger 1989;

Valuable comments on earlier drafts were provided by John Meyer, Francisco Ramirez, Claudia Buchmann, members of the Comparative Education workshop at Stanford University and the Irvine Comparative Sociology Workshop, and several anonymous reviewers. We thank Mike Zapp for providing helpful data. Authors are listed alphabetically. Direct correspondence to Jared Furuta (jkfuruta@stanford.edu) or Evan Schofer (schofer@uci.edu).

Social Forces 99(4) 1631–1657, June 2021
doi:10.1093/sf/soaa071
Advance Access publication on 28 July 2020

Kerckhoff 2001) that play a significant role in shaping outcomes such as dropout and enrollments (e.g., Heubert and Hauser 1999; Grodsky et al. 2008). The majority of countries around the world use national high-stakes educational exams (e.g., Lauwerys and Scanlon 1969; Eckstein and Noah 1992; Helms 2008). Singapore, for example, requires all students to take a series of national tests starting with the *Primary School Leaving Exam*, an exam that determines which students may continue in the system and who gains access to prestigious schools (Kwang et al. 2008). In France and many former French colonies, students take the *baccalaureate* exam at the end of the secondary school cycle (Carson 2007). In many former British colonies, students sit for the Cambridge O-level and A-level exams to gain access to higher levels of education (Raban 2008).

High-stakes exams are seen as a method of limiting educational access around the world, earning nicknames like the "iron gate" in Nepal (Davies 2015). Students spend months or even years preparing for national exams, often attending specialized test preparation schools at great expense (Baker et al. 2001; Bray 2017). Nevertheless, enrollments have grown at incredible rates worldwide, especially during more recent decades characterized by global norms of "Education for All" (Chabbott 2003). For example, secondary and tertiary enrollments in Nepal have grown ten-fold in just a few decades (World Bank 2019). How can/should we think about the "iron gate" and similar high-stakes exams, in a world of rapid educational expansion?

We argue that the relationship between high-stakes exams and educational enrollments should be understood in the context of successive waves of the global diffusion of cultural models of schooling. Exam systems reflect an early wave of diffusion, when external forces imposed or national governments imported European models of modern schooling to countries and colonial territories across the world (Benavot and Resnik 2006; Furuta 2020). In the 1950s and 1960s, advanced education was rare, and high-stakes exams represented one of several ways to channel students through the system. Exams were generally viewed as progressive and meritocratic, in comparison to alternative methods of selection that were less formal and standardized.

However, the post-1960 era saw the emergence of new global norms that emphasize educational expansion—first as a source of national economic development and later as a fundamental human right (e.g., Meyer et al. 1992; Schofer and Meyer 2005). We argue that exam regimes institutionalize older selective logics of schooling, rendering countries more resistant to later global conceptions of education that emphasize inclusivity. As a result, exams restrict enrollments *more strongly* over time, as global norms increasingly support educational expansion. High-stakes exams are thus an important and interesting case of competing institutional logics and path dependence operating across world society (Thornton and Ocasio 1999; Weir et al. 1988; Meyer et al. 1997).

Background: National High-Stakes Exams

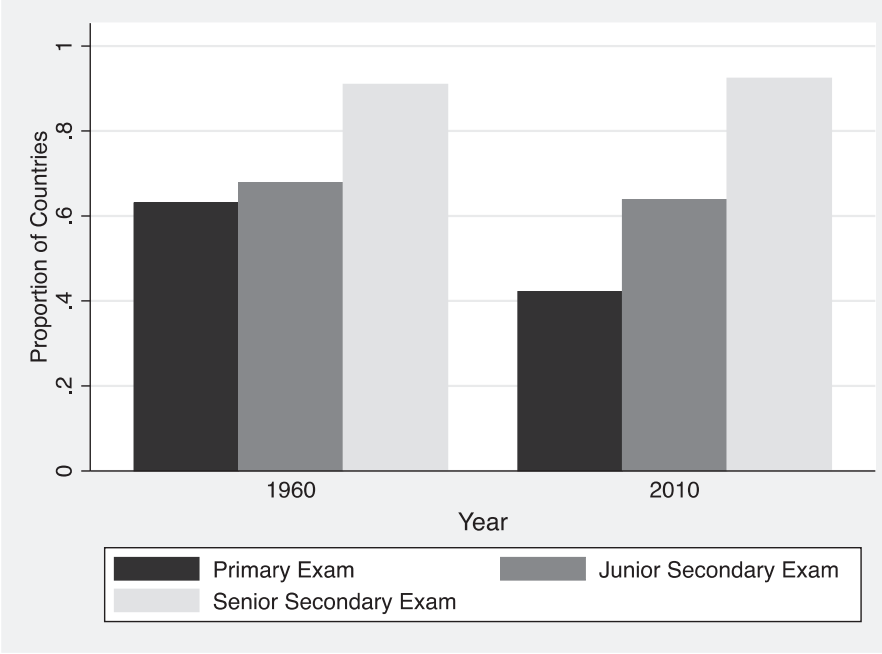
We define national high-stakes exams as tests administered to all students in a given country that determine (a) whether a student is allowed to advance in the educational system and/or (b) which of multiple sharply differentiated tracks[1] a student may enter (e.g., academic schooling vs. vocational education) (cf. Heubert and Hauser 1999). A typical example would be a national secondary school exit exam, which students must pass to earn a degree and proceed to higher education. Another common form would be an exam at the primary level that determines which students may attend academic secondary school (vs. vocational/technical secondary school). Note that our definition does not include all exams that are consequential for students. For instance, the Scholastic Aptitude Test (SAT) in the United States is sometimes referred to as a high-stakes exam because scores affect admission to selective institutions. However, students who score poorly on the SAT may nevertheless gain admission to many accredited 4-year colleges (Furuta 2017). We do not suggest that the SAT is unimportant, but this study addresses a different type of exam system—extremely common around the world—that decisively affects student advancement and/or access to academic schooling.

Furthermore, we focus on exams that have high stakes for individual students, as opposed to "assessment" exams designed to generally evaluate student performance or the quality of education in a given school, region, or nation, often for accountability purposes (cf. Ramirez et al. 2018). We investigate exam systems that are national in scope. National exam systems typically define the overall character of a society's educational system and are consequential for a wide range of outcomes (e.g., Eckstein and Noah 1992; Kellaghan and Greaney 1992).[2]

The establishment of high-stakes exams in some US states has brought renewed attention to the issue in American sociology (e.g., Heubert and Hauser 1999; Grodsky et al. 2008; Warren and Kulick 2007). While high-stakes exams in the United States are often less formidable than exams in many countries, Warren et al. (2006) nevertheless find some negative effects of exams on completion rates in US schools (cf. Warren and Jenkins 2005; Warren and Edwards 2005). Comparative education scholars have given more attention to high-stakes exams and their consequences as part of an extensive line of research that emphasizes cross-national differences in educational institutions and stratification (e.g., Turner 1960; Allmendinger 1989; Kerckhoff 2001). This work, which we draw on below, highlights the use of exams for selecting students into educational tracks, managing enrollment, providing incentives, and maintaining overall educational quality (e.g., Bol et al. 2014; Eckstein and Noah 1992; Heyneman and Ransom 1990; Kellaghan and Greaney 1992).

Figure 1 shows the proportion of countries with exams at the end of the primary, junior secondary, and senior secondary levels of schooling in 1960 and 2010. The majority of countries have high-stakes exams at

Figure 1. Proportion of countries using high-stakes exams, 1960 and 2010



some level of the educational process, most commonly at the end of the senior secondary level. Over the past 50 years, the use of high-stakes exams has declined somewhat, mainly at the primary level, as national governments made subsequent levels of schooling compulsory (see Furuta 2020).

Exams and Educational Enrollments

Exams as a constraint

Existing research generally views exams as constraints on enrollments. A large body of work argues that schools reproduce social inequalities (Collins 1979; Boudon 1973; see Breen and Jonsson 2005 for a review), and high-stakes testing is one obvious site where this may occur (e.g., Jackson and Buckner 2016). Despite claims of meritocracy, tests may reward the cultural knowledge of elite groups and provide formidable barriers to students lacking cultural capital or financial resources (Eckstein and Noah 1992; Bourdieu 1996 [1989]). For instance, affluent students often take costly test preparation courses, allowing them to outcompete poor students (Stevenson and Baker 1992; Park et al. 2016). Reproduction theories would generally expect high-stakes exams to limit enrollments, systematically

excluding lower-class and minority students from advanced forms of academic schooling.

High-stakes exams might constrain enrollments for several additional reasons. First, exams might be an organizational response to resource constraints or reflect policy preferences for emphasizing educational "standards" or "quality" over inclusiveness. Moreover, high-stakes exams represent a centralized structure that facilitates state control over access to schooling. In other words, exams may function like brakes on a car, wherein countries with exams have the capacity to slow educational growth if they wish.[3] Indeed, classic work on the US education system suggests that educational expansion in the nineteenth century occurred in part because the state lacked the ability to constrain growth (Ben-David and Zloczower 1962; Rubinson 1986). In the absence of centralized control, enrollments grew without much regard for capacity or school quality.

Whether the argument centers on elite reproduction or policy preferences, the empirical expectation is the same:

Proposition 1: National high-stakes exams are an organizational constraint that limits educational enrollments. □

Exams as meritocratic selection and incentive

By contrast, some scholars argue that exams can boost educational quality without unduly limiting participation. For example, students who are unable to pass a national high-stakes exam may be those who were likely to drop out anyway or who have little chance of succeeding in higher levels of the educational system (e.g., Jacob 2001). Moreover, proponents of exams argue that they improve the overall quality of schooling in a society by motivating students and holding teachers and administrators accountable (Bishop 1997; Heyneman and Ransom 1990; Raymond et al. 2003). Indeed, the use of centralized exams often prompts curricular standardization (Allmendinger 1989), potentially benefiting students who might otherwise be marginalized or receive less attention in lower school tracks (e.g., Toenjes et al. 2002; cf. Bol et al. 2014). If so, examination regimes might help weaker-performing students succeed, maintaining or perhaps even boosting enrollments. Thus:

Proposition 2: High-stakes national exams may not reduce educational enrollments and could even increase enrollments. □

High-Stakes Exams in Historical and Institutional Context

We seek to reframe discussions about high-stakes exams within a broader global and historical context. The meaning and consequences of a particular educational policy, like high-stakes tests, depend on the broader institutional context in which the policy is embedded. Comparativists develop this point to explore variabilities among national systems (e.g., Allmendinger 1989; Kerckhoff 2001). We theorize that global historical shifts—chronicled by neo-institutional scholars of education—transform the meaning and consequences of high-stakes tests.

We draw on institutional theories of education (Meyer 1977; Baker 2014), comparative studies of education systems (e.g., Turner 1960; Allmendinger 1989; Kerckhoff 2001), and historical institutionalism (Thelen and Steinmo 1992; see Amenta and Ramsey 2010 for a review). The institutional tradition views education as a foundational component of modern society and culture: defining key categories of knowledge, personnel, and expertise, effectively constructing "schooled" societies (Meyer 1977; Baker 2014; Schofer and Meyer 2020). Institutional scholars argue that the global growth of schooling is driven by norms, policy models, and ideologies that are propagated across the international community, as opposed to being driven purely by local factors such as industrialization, demand for skilled workers, or elite interests (Meyer et al. 1992). Thus, a major predictor of enrollment growth is country ties to international organizations that propagate norms in favor of educational expansion (Schafer 1999; Schofer and Meyer 2005).

We make several arguments, which collectively posit that the effects of exams on enrollments may change over time. (1) The institution of education has shifted from a limited and specialized enterprise in the early twentieth century to an expansive human right by the end of the century. (2) At different points in time, the dominant institutional models of education tended to diffuse across the world. (3) Countries that adopted testing in the early period effectively institutionalize policies that render them more resistant to later norms favoring educational inclusion.

The Great Educational Transformation: From Limited to Mass Schooling

Testing in a world of limited schooling

Mass schooling is largely a twentieth-century phenomenon. In the nineteenth century, opponents of educational expansion questioned whether peasants were educable and warned that mass schooling would lead to rising expectations and social disruption (Maynes 1985). The early twentieth century saw an emergent consensus across Europe that mass education (of white males) was beneficial, but this generally meant elementary school. Secondary and tertiary schooling did not start growing rapidly until decades later. In 1950, the country-average secondary school enrollment ratio in the world was approximately 12 percent, while tertiary ratios averaged only 2 percent (Barro and Lee 2015).[4]

Exams as scientific selection

In the early twentieth century, most assumed that few people would pursue higher levels of schooling. The only question was *how* to manage the process of sorting and selecting students. Societies, and particular localities or schools, pursued a range of strategies: streaming students into different terminal and nonacademic school "tracks," establishing quotas or *numerus clausus*, brusque forms of exclusion based on students' ascriptive status characteristics or political views, and so on (Windolf 1997; Ringer 1979).

The United States, with its long history of decentralization, never established an educational ministry or national selection system, and, consequently, educational selection remained locally heterogeneous. In contrast, most other "modern" countries established government ministries tasked with building and managing an overall national education system (Ramirez and Ventresca 1992). This process generated rationalization and "governmentality," through which state bureaucracies imposed order and control over society (Weber 1978 [1922]; Foucault 2009). Ministries began to plan and forecast enrollments, and often managed the number of graduates in different fields.

In this early period, educational ministries often turned to high-stakes exams as a "rational" and efficient way to select students. At the time, many viewed exams as quite progressive: putatively impartial, scientific, and meritocratic[5] compared to selection criteria such as teacher evaluations or ascriptive status characteristics (Benavot and Resnik 2006; Windolf 1997). By mid-century, the European model of state-managed schooling spread across much of the world (Ramirez and Rubinson 1979), and high-stakes testing regimes often traveled with this (Coleman 1965; Raban 2008). In some cases, independent countries emulated the policies of successful industrialized powers like France and Great Britain (Furuta 2020). In other cases, colonial powers imposed Western-style schooling and/or provided aid and organizational templates for the establishment of school systems after colonies gained independence (e.g., Watson 1982).

Education-for-development and Education for All

The 1960s saw the rise of new educational policy discourses favoring expanded enrollments, which spread rapidly across industrialized countries via key international organizations such as UNESCO and the World Bank. Drawing on new economic ideas, such as human capital theory, UNESCO called for the expansion of education to accelerate industrialization and economic growth (Mundy 1999; Chabbott 2003). A new consensus emerged, leading to massive allocations of state resources and international development aid toward educational expansion—and enrollments accelerated worldwide (Drori et al. 2003).

The education-for-development logic of the 1960s fit fairly well with prior national testing regimes. Many countries turned to *manpower planning*, tailoring educational tracks, and enrollments to meet the perceived needs of industrializing economies. States viewed high-stakes exams as a meritocratic way to allocate students across these systems. However, as UNESCO and other international organizations began to recommend expanding mass education beyond primary school, as part of efforts to boost human capital, expansionary logics began to conflict with early high-stakes testing. A few countries abandoned primary-level high-stakes testing in this era (Furuta 2020).

The 1980s and 1990s saw another wave of educational policy discourse, reframing education as a fundamental human right. This ideology quickly gained support from UNESCO and others, leading to major international declarations

in Jomtien, Thailand (1990), and Dakar, Senegal (2000), that focused on *Education for All* (EFA). The EFA movement stressed educational inclusion, not merely to meet the needs of the labor force but as a means to improve society and human life broadly and as an end unto itself.[6]

National institutionalization, competing logics, and policy feedback

We argue that national systems of high-stakes testing institutionalize older logics of limited and selective schooling—slowing national conformity to the newer logic of Education for All. As comparative sociologists of education emphasize, features of national education systems become deeply entrenched, forming durable configurations (e.g., Busemeyer 2015; DiPrete et al. 2017; Kerckhoff 2001). Such historically evolved systems have path-dependent consequences, which historical institutionalists refer to as *policy feedback*: Prior policy choices, once institutionalized, create organizational constraints and reify social attitudes, which can then constrain future possibilities (Thelen and Steinmo 1992; Weir et al. 1988). One can think of this in terms of alternative institutional logics (Thornton and Ocasio 1999) that conflict: countries that adopt high-stakes testing regimes effectively "lock in" selective logics, slowing the adoption and implementation of newer educational logics that support enrollment expansion.

We suggest several ways that a prior commitment to testing might slow enrollment expansion in the Education for All era. First, high-stakes exams embody and reify an older cultural understanding of education as a selective enterprise: students are presumed to be qualitatively different in their capacities, and some simply do not have the requisite academic abilities required for advanced stages of schooling (see Kandel 1930). Such beliefs conflict with the more recent logic of Education for All and thus create potential for resistance to enrollment expansion.

In addition, high-stakes exam systems are embedded in a broader set of national policy regimes that shape the overall selectivity of the system, including other educational institutions (e.g., tracking, compulsory schooling laws) and more general "skills" regimes (e.g., school-to-work transitions, linkages between state educational institutions and firms) (Hall and Soskice 2001; DiPrete et al. 2017). For example, school systems often use high-stakes exams to sort students into sharply differentiated school "tracks," in which comparative research has shown to substantially increase educational inequalities (e.g., Allmendinger 1989; Bol and van de Werfhorst 2013). Historically, the evolution of these national policy regimes has led to competing institutional narratives that shape educational expansion (e.g., ideals of educational "opportunity" in the United States vs. welfare state "security" in Germany; see Heidenheimer 1981). Thus, even if policymakers design high-stakes exams to sort and select students through formally fair and meritocratic processes, they become an integral component of a broader set of state institutions that may reflect more selective logics of education.

Furthermore, exams require large ministerial bureaucracies to estimate labor force needs, devise future enrollment plans, and so on. Not only does this provide

the organizational capacity to limit school expansion, it creates information to justify such decisions and legitimates the very idea that access should be limited. Exam regimes tend to institutionalize "standards" or "quality" as a central focus of ministerial planning (Allmendinger 1989; Bol et al. 2014) and produce a constituency of ministerial officials and specific divisions within ministries oriented toward those ends.

Finally, exam regimes extend far beyond the state's organizational machinery. Typically, the entire societal elite is a product of the exam system, and thus the societal status hierarchy becomes anchored in older visions of education as a highly selective enterprise. In countries like Japan and South Korea, for example, where admission to the country's most prestigious universities is typically decided entirely by a student's entrance exam score, most business, government, and professional elites are graduates of these top universities (Rohlen 1983; Seth 2002). Rohlen (1983, p. 90) notes ironically that "in [Japan's] Ministry of Education — explicitly assigned the task of disassembling the elite university influence over exams and jobs — sixteen out of eighteen of the top positions were filled by Tokyo University graduates." It is not surprising that elites produced by exam systems tend to believe in the virtues of exams and selectivity and tend to resist expansionary reforms.

In sum, national testing regimes institutionalize an early twentieth century educational logic of selection and thus serve as a source of inertia or resistance to recent global norms favoring educational expansion. Countries without testing regimes are more easily swept along by recent norms of Education for All. Thus, we advance the following propositions:

Proposition 3: National high-stakes testing regimes are associated with lower enrollments in the Education for All era (i.e., after 1990 and/or as global pro-education organizations and norms become entrenched). □

Proposition 4: National high-stakes exam regimes dampen the expansionary effect of international pro-education organizations and norms on national societies. □

Data and Methods

We explore these arguments using cross-national data on educational testing and educational enrollments for 142 countries from 1960 to 2010. The unit of analysis is the country-year. Our dataset, which includes 4,841 country-years, is unbalanced because some countries were not independent in early years and due to missing data.

Dependent Variables

Secondary enrollments

We measure secondary school participation as the gross secondary enrollment ratio from the *World Bank Development Indicators* dataset (World Bank 2019).

The gross enrollment ratio is simply the number of enrolled students divided by the number of people in the official age group for secondary school in each country.[7] For instance, if a country's age for secondary attendance is 11 to 18 years, the gross enrollment ratio would be the number of secondary students divided by the total population aged 11–18. We also examined "net" enrollment ratios, which only include enrolled students who are of the official school age in the numerator (i.e., excluding overage and underage enrollees). Results were similar; however, sample sizes were much smaller, so we opted to use the gross enrollments measure.

Tertiary enrollments

We use the gross tertiary enrollment ratio from the *World Bank Development Indicators* (World Bank 2019). Again, this is simply the number of students enrolled in tertiary schooling, divided by the number of individuals in the relevant population age group.

Independent Variables

High-stakes testing data

Our analyses include dichotomous variables that indicate a country's use of *high-stakes exams* at the primary, junior secondary, and senior secondary levels of schooling.

We constructed a wholly new cross-national dataset on high-stakes testing, coded from several sources: international education encyclopedias, UNESCO surveys on the organization of education in different countries around the world, and formal reports from UNESCO's international conferences on education. We include a full list of sources in Supplementary Appendix I.

Our coding of *high-stakes exams* follows the definition discussed above: nationally administered exams that determine whether a student is allowed to continue in the education system and/or which among sharply differentiated tracks a student is allowed enter (e.g., academic vs. vocational high school). For example, all students in The Gambia in the early 2000's took the *Common Entrance Exam* at the end of sixth grade, which determined placement in the junior secondary system. The system allowed high scorers to enroll in one of seven available academic high schools, while their lower-scoring classmates could enter the vocational junior secondary system or drop out. To continue in the education system and qualify for university entrance, students must have passed the *ordinary-level* exam at the close of academic junior secondary education and the *advanced-level* exam at the close of academic senior secondary education. In our dataset, we coded The Gambia as having three high-stakes exams (UNESCO 2006).

We first identified the presence of high-stakes exams in each decade from 1960 to 2010. When changes were observed (e.g., an exam existed in the 1970s but not in the 1980s), we drew on our sources to identify the year that a change took place.[8] This allowed us to construct an annual time series dataset. In

the coding process, we validated the accuracy of our data with other primary sources, secondary literature, or other data collected for similar projects (e.g., Bol and van de Werfhorst 2011). For example, data collected for the 1990s relied on overlapping information from the *International Encyclopedia of Education* (Husen and Postlethwaite 1985), the *Encyclopedia of Higher Education* (Clark and Neave 1992), and 130 UNESCO conference reports from the *International Conference on Education*.

Gross domestic product per capita (logged)

Economic development, as indicated by real GDP per 10,000 capita, comes from the *Penn World Tables* (Feenstra et al. 2015). Prior research indicates that economic development may prompt educational expansion, although empirical findings are mixed (e.g., Windolf 1997; Schofer and Meyer 2005).

Democracy

We capture a country's level of democracy using data from the *Polity IV Project* (Marshall et al. 2014), where a "−10" indicates a "strongly autocratic" regime and a "+10" indicates a "strongly democratic" regime. Some studies suggest that democratic regimes have higher enrollments (e.g., Benavot 1996), which may reflect societal policy preferences for inclusive schooling.

International nongovernmental organization memberships (INGOs) (logged)

This is a standard indicator of a country's linkages to the international community or "world society." Scholars have shown that international organizational ties serve as a useful proxy for a country's exposure to global norms and pressures (Boli and Thomas 1999), and prior empirical work observes that INGO memberships tend to predict enrollment growth (Schafer 1999; Schofer and Meyer 2005). The measure is a country's number of citizen memberships in INGO's (logged to reduce skewness), based on the *Yearbook of International Organizations* (Union of International Associations 1960–2010).

Global Education for All norms

We measure the global shift toward "Education for All" (EFA) norms in two ways. First, we employ a post-1990 time dummy variable, which contrasts early decades with the period following the Jomtien conference's Education for All declarations.[9] Second, we constructed an annual time-varying index that reflects the number of international pro-education organizations in the world, logged (Ramirez et al. 2016); the number of pro-education conferences in the world, logged (Zapp and Ramirez 2018); a post-1990 dummy variable indicating years after the Jomtien conference; and the number of international nongovernmental organizations in the world (a measure of global cultural influence, described above) (Union of International Associations 1960–2010). We constructed an index by taking the z-score of each component measure and summing them together (Cronbach's alpha = 0.78).

Table 1. Descriptive Statistics

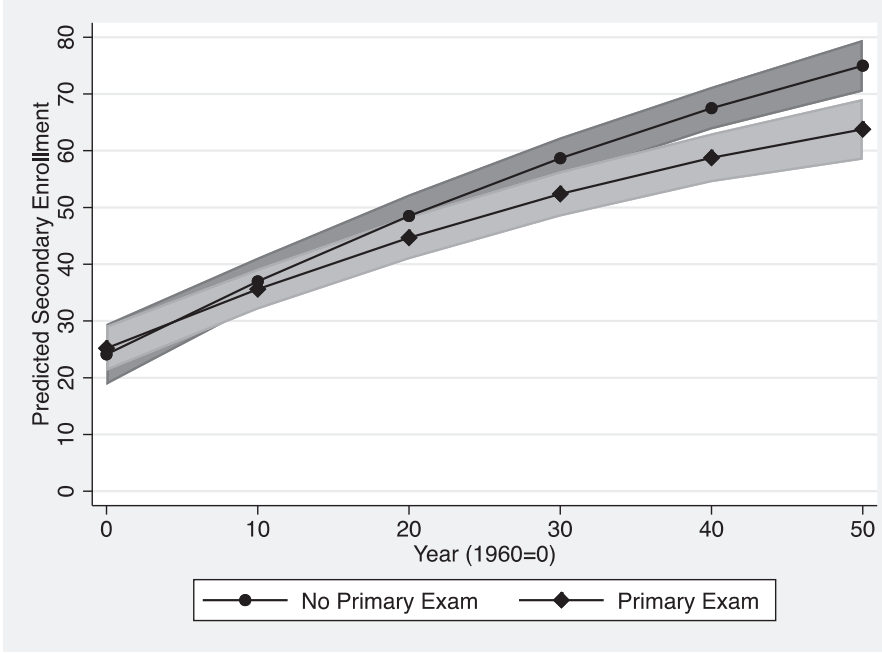
Variable	Obs.	Mean	Std. dev.	Min	Max
Primary enrollment	4,841	91.81	24.64	7	165.65
Secondary enrollment	4,841	52.39	34.13	1	161.02
Tertiary enrollment	4,841	16.06	19.04	0	106.64
Primary exam	4,841	0.45	0.5	0	1
Jr. secondary exam	4,841	0.61	0.49	0	1
Sr. secondary exam	4,841	0.89	0.31	0	1
GDP per capita (log)	4,841	8.69	1.24	5.24	11.96
Democracy	4,841	1.44	7.55	-10	10
INGO memberships (log)	4,841	6.06	1.06	2.2	8.36
Late period dummy	4,841	0.66	0.47	0	1
World pro-education orgs and norm index	4,841	0.82	2.66	-4.65	4.81

Table 1 presents descriptive statistics of all variables. Supplementary Appendix G presents a correlation matrix, and Supplementary Appendix H lists the countries included in our main analyses. To reduce missing data, we used linear interpolation to fill gaps under 5 years in length. Our independent variables, such as GDP, tend to change smoothly over time, so this does not seem likely to add a great deal of bias (interpolation does not alter the findings). We used listwise deletion to ensure a consistent sample across all analyses.

Methods

We explore the relationship between national high-stakes exams and school enrollments using annual country-level data from 1960 to 2010. We present panel regression models with country random effects and cluster-robust standard errors (Wooldridge 2010; Baltagi 2013). Random effects are appropriate given that our core variable of interest (exams) change infrequently, and thus most variation is cross-sectional. Country fixed effects, by contrast, would focus only on temporal change within the few cases in which exams were added or removed. That said, we observe similar results with both random and fixed effects models, as well as other common panel regression strategies such as OLS regression with panel-corrected standard errors (not presented; available from the authors upon request).[10] We also considered clustering by time in addition to country ("two-way" random effects), and results were similar (available upon request). We lagged all independent variables by 1 year. We found similar results using longer lag times (e.g., 3 or 5 years). Finally, we examined Cook's D and partial regression plots to identify influential cases. Removing the small number of moderate outliers did not change the results; we decided to retain these cases in the sample.

Figure 2. Predicted secondary enrollment ratio, for countries with and without high-stakes exams. (Time = 0 refers to 1960). Note: Predicted enrollments estimated from our base model, with additional controls for time, time-squared, and an interaction of exam * time. See Supplementary Appendix A

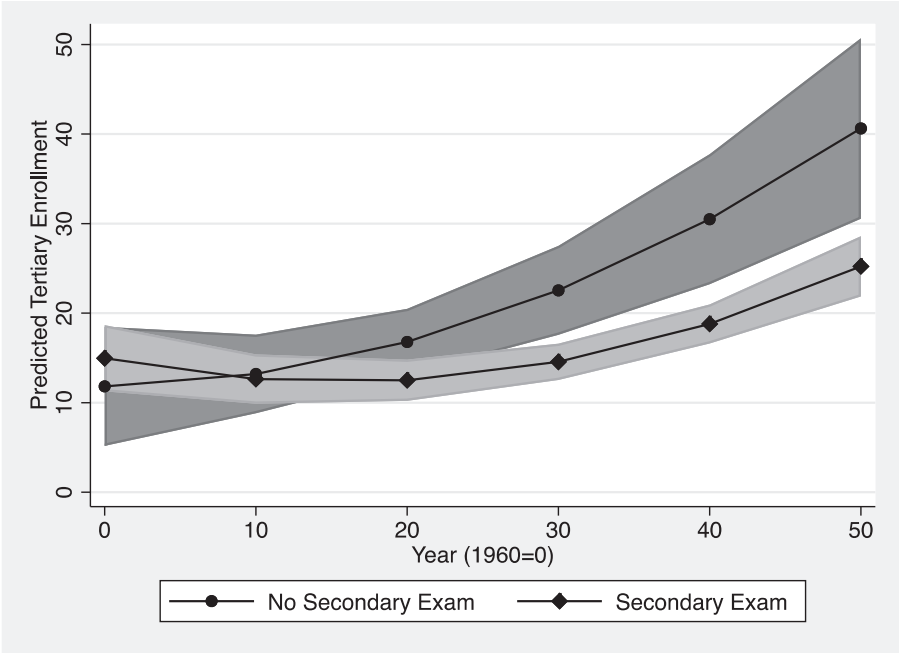


Results

We begin with a visualization of educational enrollments for countries with and without exams to provide a sense of the historical trends. Figures 2 and 3 present predicted secondary and tertiary enrollments based on results of a panel regression model that includes controls for GDP per capita, democracy score, INGO memberships, polynomial time, and an interaction between exams and time (results presented in Supplementary Appendix A, Models 3 and 4). The purpose is to highlight the time trend of enrollments and net of controls. In the early years, countries with national exams have similar enrollments to those without exams. In recent years, countries with high-stakes exams have fewer secondary and tertiary students by a substantial margin—on the order of 10 percentage points. These trends are consistent with our argument that exam systems have little effect on enrollments in the era of limited schooling but become increasingly consequential as the world shifts toward Education for All norms.

Table 2 presents panel regression models examining the relationship between high-stakes exams and secondary and tertiary enrollments. We observe consistent negative associations between high-stakes exams and educational expansion

Figure 3. Predicted tertiary enrollment ratio, for countries with and without high-stakes exams. (Time = 0 refers to 1960). Note: Predicted enrollments estimated from our base model, with additional controls for time, time-squared, and an interaction of exam * time. See Supplementary Appendix A



and net of controls. Countries with exams at the primary level have secondary enrollments that are significantly lower than countries without exams. Likewise, senior secondary exams are associated with lower tertiary enrollments.[11] The coefficients are substantively large: Gross enrollment ratios are roughly 6 to 7 percentage points lower in countries with exams. In sum, the results in table 2 are consistent with the proposition that high-stakes exams act as organizational barriers that constrain educational expansion.

Table 3 explores the idea that the wider global institutional environment, which increasingly emphasizes norms of Education for All, shapes the effects of high-stakes exams on educational enrollments. We interact high-stakes exams with the following measures: (a) historical period (before vs. after the first worldwide Education for All conference in 1990), (b) an index that reflects global norms supporting educational expansion, and (c) a country's INGO memberships, which capture the extent to which a country is embedded in global norms. Table 3 focuses on the most relevant exam for a given level of the educational process: primary exams for models of secondary enrollment and senior secondary exams for models of tertiary enrollment.

Results in table 3 are consistent with our core argument. In table 3, Models 1 and 2, there is a large and significant negative interaction between exams

Table 2. The Effect of Educational Tests on Secondary and Tertiary Enrollments

	Secondary enrollment				Tertiary enrollment			
Primary enrollment	0.06	0.07^{+}	0.06	0.07^{+}				
	(0.042)	(0.042)	(0.043)	(0.041)				
Secondary enrollment					0.34***	0.34***	0.34***	0.34***
					(0.053)	(0.053)	(0.048)	(0.049)
GDP per capita (log)	13.48***	14.45***	14.48***	13.51***	7.50***	7.60***	7.60***	7.52***
	(1.810)	(1.860)	(1.881)	(1.771)	(1.819)	(1.802)	(1.810)	(1.807)
Democracy	0.17	0.16	0.16	0.15	0.12	0.13	0.10	0.11
	(0.142)	(0.138)	(0.145)	(0.139)	(0.101)	(0.101)	(0.099)	(0.099)
INGO memberships	14.61***	14.78***	14.86***	14.54***	2.13*	2.06*	2.16*	2.10*
	(1.479)	(1.474)	(1.472)	(1.433)	(0.939)	(0.940)	(0.901)	(0.908)
Primary exam	−7.09***			−6.29***	−0.50			−0.83
	(1.787)			(1.891)	(1.682)			(1.677)
Jr. secondary exam		−5.85*		−3.18		1.00		2.41
		(2.684)		(2.757)		(1.772)		(1.734)
Sr. secondary exam			−6.25	−4.23			−6.78**	−7.64**
			(7.183)	(7.152)			(2.354)	(2.654)
Constant	−152.95***	−162.41***	−160.02***	−148.02***	−78.17***	−79.66***	−73.08***	−72.48***
	(13.691)	(13.541)	(14.786)	(14.247)	(16.643)	(16.485)	(16.733)	(16.755)
Observations	4,841	4,841	4,841	4,841	4,841	4,841	4,841	4,841
Countries	142	142	142	142	142	142	142	142

Robust cluster standard errors in parentheses.
***$P < .001$, **$P < .01$, *$P < .05$, $^{+}P < .10$

Table 3. The Effect of High-Stakes Educational Tests on Secondary and Tertiary Enrollment: Interaction Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Secondary	Tertiary	Secondary	Tertiary	Secondary	Tertiary
Primary enrollment	0.09*		0.08*		0.09*	
	(0.039)		(0.035)		(0.042)	
Secondary enrollment		0.33***		0.25***		0.32***
		(0.046)		(0.044)		(0.043)
GDP per capita (log)	12.76***	7.59***	11.86***	8.65***	12.32***	7.65***
	(1.748)	(1.867)	(1.584)	(1.932)	(1.813)	(1.846)
Democracy	0.12	0.08	−0.06	0.02	0.13	0.11
	(0.128)	(0.098)	(0.129)	(0.101)	(0.141)	(0.095)
INGO membership	10.87***	2.47*	3.00+	−1.69	16.98***	8.34**
	(1.754)	(1.007)	(1.786)	(1.180)	(1.991)	(2.706)
Exam	−2.20	−1.33	−5.62***	−5.05*	−4.32+	0.49
	(2.327)	(2.567)	(1.636)	(2.121)	(2.419)	(3.834)
Late period (post-1990)	10.52***	7.09*				
	(1.647)	(2.922)				
Exam * late period	−7.86**	−8.02*				
	(2.401)	(3.194)				
Global pro-education orgs and norms			3.72***	2.73***		
			(0.409)	(0.667)		

Continued

Table 3. Continue

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Secondary	Tertiary	Secondary	Tertiary	Secondary	Tertiary
Exam ∗ global orgs and norms			−1.23**	−1.52*		
			(0.440)	(0.688)		
Exam ∗ INGO memberships					−4.41*	−6.50*
					(1.825)	(2.748)
Constant	−2.79***	−75.34***	−59.15***	−66.80***	−69.20***	−68.05***
	(0.658)	(14.414)	(13.114)	(15.692)	(14.740)	(15.219)
Observations	4,841	4,841	4,841	4,841	4,841	4,841
Countries	142	142	142	142	142	142

Notes: Robust cluster standard errors in parentheses. "Exam" refers to the relevant exam for the outcome variable: primary exams for models of secondary enrollment and senior secondary exams for models of tertiary enrollment.
$^{***}P < .001$, $^{**}P < .01$, $^{*}P < .05$, $^{+}P < .10$

and the period after 1990, when global norms of Education for All became established. (A linear time interaction shows a similar pattern; see Supplementary Appendix A, models 1 and 2.) The coefficient for exams, which indicates the effect size prior to 1990 due to the inclusion of an interaction, is small and not significant. By contrast, the negative and significant interaction indicates that tests are associated with lower enrollments after 1990.

Models 3 and 4 in table 3 include a more detailed index of international pro-education conferences and organizations. Again, we see a strong negative interaction effect, wherein exams are a stronger constraint on enrollments as Education for All becomes globally dominant. Finally, results in Models 5 and 6 suggest that the influence of global norms on national enrollments—operationalized here as ties to international organizations—is weaker in countries with high-stakes exam regimes. This fits with our argument that prior institutionalization of selective logics renders countries more resistant to global pro-education norms and pressures. Finally, the coefficients of the post-1990 dummy, the world education index, and INGO memberships are also all positive and statistically significant, again consistent with our argument that global educational pressures and norms drive enrollment expansion, particularly in countries without national exams.

Corollary analyses: additional controls and interactions

We conducted extensive corollary analyses, both to address potential sources of omitted variable bias and to explore additional arguments of interest. We summarize the key points here (corollary analyses are presented in Supplementary Appendices A–F). We controlled for additional measures such as historical time (year), the size of the service sector (value added as % of GDP, an indicator of demand for skilled workers) (World Bank 2019), size of the welfare state measured as social transfers as % of GDP (World Bank 2019), unemployment based on ILO estimates (World Bank 2019), economic inequality (Gini score) (Solt 2014), economic growth (% change in GDP), economic crisis (a dummy reflecting >10% annual decline in GDP), and other factors. These controls did not alter our main findings.

We also considered factors that might intensify or dampen exam effects (i.e., interactions). First, countries that lack fiscal resources to invest in their education systems might stringently select a small number of high-performing students in order to maximize the returns on their investments. Thus, a country's GDP per capita might interact negatively with exams. Second, countries with higher levels of economic inequality may be those in which powerful elites are able to reproduce their advantages via the education system, causing restrictive effects of exams to be larger. Third, the effect of exams may be smaller in countries with extensive privatization of schooling, which may allow individuals to circumvent the effects of national exams. Fourth, a country's former colonial history might moderate the effect of high-stakes exams on enrollments. For example, former French and British colonies may use high-stakes exams to legitimate and maintain the selective and elite-driven educational systems that

were common in colonial contexts (cf. Watson 1982; Furuta 2020). Results are presented in Supplementary Appendices D and E.[12]

Finally, Supplementary Appendix F includes the interaction of all variables with exams.[13] The interaction models reveal some intriguing findings, which warrant future research. First, the interaction of exams with a country's GDP per capita is negative, indicating that exams are a more formidable constraint in affluent countries. This runs counter to the common sense expectation that poor countries impose stringent exams because they lack resources to expand schooling. Rather, it is among affluent countries—that could presumably afford to expand schooling—where exams are most consequential in constraining growth. One interpretation is that affluence provides the resources for hyper-expansion in the absence of organizational constraints (Ben-David and Zloczower 1962; Rubinson 1986). Or, the effect might be spurious: exam regimes in affluent countries tend to be old and deeply entrenched and thus may strongly institutionalize logics of educational selection. Also, the interaction between secondary enrollment and secondary exams is negative in analyses of tertiary enrollments. This is consistent with conventional reproduction imagery: elites seek to limit expansion at higher levels of the system as lower levels fill up. Or, this may reflect the use of exams to manage enrollments in countries with limited tertiary capacity. In any case, our findings regarding Propositions 1–4 are generally robust.[14]

Discussion and Conclusion

Drawing on a newly created cross-national dataset of high-stakes exams in 142 countries from 1960 to 2010, we find evidence that exam systems are associated with lower enrollments at the secondary and tertiary levels.[15] This fits with reproduction theories, as well as the conventional understanding that exams represent a formidable hurdle for students.

Upon closer examination, however, the constraining effect of testing regimes on enrollments is most evident in recent decades. Drawing on institutional theories of education, we argue that high-stakes exams were less consequential in the mid-twentieth century because the institution of education was very different. Prior to the advent of the "schooled society" (Baker 2014), few jobs required advanced education, and few people pursued schooling. In that context, exams were often viewed as a progressive alternative to selection on ascriptive characteristics, subjective personal evaluations, or "old boy" networks. Countries with high-stakes exams do not appear very different from those without exams in the early years of our study; enrollments were low everywhere.

Over time, international organizations and governments reimagined education as a source of human capital and later a human right, and enrollments grew worldwide. Countries with centralized exam regimes possessed institutionalized structures that could resist this inclusive view of schooling: Extensive bureaucracies devoted to managing and limiting school access, and elites that were the product of exam-based selection. Soaring international visions of Education for

All run up against older logics of manpower planning and elite selection.[16] In systems where student test scores and labor force needs are carefully monitored, it becomes easier for ministerial officials or others to raise concerns about the erosion of educational standards or "wasteful" overeducation.

Our discussion extends institutional theories of policy diffusion. Prior research focuses on single waves of policy diffusion. We point out that successive waves of educational diffusion create tensions, as earlier policy adoption, or the failure to adopt, may affect later waves. Thus, historical institutional ideas of path dependence and policy feedback—mainly studied in particular national contexts—provide leverage to explain dynamics of global diffusion. Finally, the case of high-stakes exams exemplifies a key insight about institutional change: New norms or institutions do not simply and, entirely, displace a prior set of institutions. Instead, new norms and conceptions of education wash over previously institutionalized policy regimes, sometimes creating institutions with conflicting logics that are "layered" on top of each other (Mahoney and Thelen 2010).

Conventional discussions of high-stakes exams are often polarized around the question of whether exams represent malign sites of class reproduction or meritocratic incentives. It is difficult to assess whether national exam systems are substantively more or less meritocratic than systems that determine educational access based on criteria like teacher evaluations and grades. Elites adapt in either case: for instance, via test preparation and "cram schools" in educational systems with exams and via carefully curated grades, resumes, and teacher recommendations in systems that do not feature exams. Yet, our study suggests that the elimination of exams may lead to more rapid enrollment expansion—at least during the current era of inclusive educational norms.

Beyond the issue of enrollments, exam regimes may shape inequality more broadly. Exams legitimate education-based social distinctions, which Domina et al. (2017) refer to as categorical inequalities. The ideology underpinning exam regimes may seep into society, influencing understandings of intelligence, ability, and merit. Future research could explore the relationship between national exam regimes and individual attitudes as well as aggregate societal inequalities.

Exam regimes, finally, are instances of *institutionalized state policy*. Institutional theories help explain their spread (Furuta 2020), and we suggest that historical institutionalism also helps explain their consequences over time. Exam-based selection regimes institutionalize a particular model of education, focusing on selectivity. This understanding of schooling may lead to continued educational restriction, even among countries that have sufficient resources to support educational expansion. That said, we should not overstate the inertial effects of exam systems on restricting access to education. The Education for All movement, and related ideologies, continue to sweep the globe. Rapid educational expansion has occurred everywhere, and in some cases global expansionary norms have resulted in the elimination of exam systems (Furuta 2020). National policy regimes may slow the impact of global pressures, but generally do not stop them (Hironaka 2014).

Notes

1. Here, we focus on tracks that place students into different types of schooling, typically in entirely different facilities, as opposed to ability grouping in the United States (e.g., regular vs. honors classes).
2. Also, it is generally not feasible to collect historical data on school-based exams across several decades for a large number of countries.
3. Most countries supported rapid enrollment expansion throughout the period of our study, but there are exceptions. For example, communist countries in Eastern Europe restricted higher education in the 1970s and 1980s (Baker et al. 2007).
4. One might imagine that enrollments were low in the mid-twentieth century because schools were brutally selective, but that is not the case. American Ivy League colleges and other elite schools like Oxford accepted the majority of white male applicants prior to World War II (Soares 1999; Bingham 1956). Rather, the public did not see advanced schooling as especially useful: only a few professional careers required higher education, and many of those were not well-paying (e.g., clergy, teacher, scholar). Employers typically hired individuals for occupations like managers based on personal connections or substantive experience. Especially in the United States, college was a consumption good where one might become cultured, enjoy social life and athletics, and build social networks (e.g., Karabel 2005).
5. Obviously, they may not have been meritocratic in practice.
6. The Dakar Framework for Action states "[A]ll children, young people and adults have the human right to benefit from an education that will meet their basic learning needs in the best and fullest sense of the term, an education that includes learning to know, to do, to live together and to be. It is an education geared to tapping each individual's talents and potential, and developing learners' personalities, so that they can improve their lives and transform their societies" (World Education Forum 2000).
7. Note that gross enrollments may routinely exceed 100 percent (see World Bank 2020 for a discussion).
8. Exams change infrequently. If exams were present at the start and end of a decade, we coded the exam system as being in place during intervening years. In a few cases, we were unable to determine the exact year of the change in exam systems, or the timing of implementation was unclear. In these cases, the value of a given country's exam variable prior to the change in the use of the exam was assigned for the full decade it was first identified for (e.g., from 1960 to 1969, 1970 to 1979, and so on), and the change in the variable's value was assigned at the start of the subsequent full decade.
9. In fact, the EFA movement had substantial momentum before the 1990 conference. To address this, we also examined alternative cutoffs, such as 1985. Results were consistent.
10. Hausman tests found no significant difference between fixed and random effects in some of our models. In other models, we observed a significant difference, but differences were substantively small. In analyses involving

large samples (in our case, nearly 5,000 country-years), the Hausman test frequently rejects the null, even if there is no practical difference between fixed and random effects—and thus no indication of substantial bias in the random effects results (see Wooldridge 2010).

11. Exams at lower levels of the system do not affect tertiary enrollments, which makes sense given that we control for secondary school enrollments (which presumably mediate the effects of early exams).
12. Continuous measures are mean-centered for interaction models to improve interpretability of main effects. Sample sizes vary due to limited availability of some measures.
13. VIF values were very high in the fully interactive model, indicating multicollinearity; results should be interpreted with caution.
14. The interaction of INGO membership and exams loses significance in one instance, Supplementary Appendix F, Model 10, the fully interactive model of tertiary enrollments. This may be the result of excessive collinearity in the fully interactive model, noted above.
15. Our study is observational, which poses challenges when it comes to drawing causal inferences. It is possible, for instance, that exam effects in later years reflect unobserved policy preferences that are correlated with exam regimes. We attempt to control for relevant factors, like democracy, which may reflect orientations toward inclusive schooling, but we cannot rule out the possibility of spurious effects.
16. France's Minister of Public Instruction captures this image of high-stakes exams: "Our task, in fact, is constantly to create an *elite* . . . at a time when other forces, less noble, seek to establish themselves [i.e., democracy], the educational system has the glorious privilege of maintaining the prestige of the spirit, the sovereignty and independence of the idea. [The education system] has the duty of courageously teaching democracies that no greater danger menaces them than misunderstanding the role of the *elite*, or, in other words, the danger of leveling downward. Our system by examinations and by competitions maintains the most justifiable method of recruiting" (Minister of Public Instruction of France in 1927; quoted from Kandel 1930, p. 226).

About the Authors

Jared Furuta is a postdoctoral researcher at Stanford University, where he received his PhD in sociology. His research focuses on macro-level cultural processes that shape national educational institutions, changes in conceptions of merit in US higher education, and institutional tensions in rationalized models of liberal individualism. His recent work has appeared in *Sociology of Education*.

Evan Schofer is a professor of sociology at the University of California, Irvine. His research examines global expansions of NGOs, environmentalism, and education, as well as comparative differences in political participation. Much of his work seeks to develop and extend world society theory, to better understand global patterns of social change. He received his PhD in sociology from Stanford University.

Shawn Wick is an associate professor of sociology at Central College in Pella, IA. His research examines organizational legitimacy in the field of international development, comparative-historical variation in national education systems, and community development in practice. He received his PhD in sociology from the University of Minnesota.

Supplementary Material

Supplementary material is available at *Social Forces* online, http://sf.oxfordjournals.org/.

References

Allmendinger, Jutta. 1989. "Educational Systems and Labor Market Outcomes." *European Sociological Review* 5(3): 231–250.

Amenta, Edwin and Kelly Ramsey 2010. "Institutional Theory". In *Handbook of Politics—State and Society in Global Perspective*, edited by Leicht, Kevin, Jenkins, Craig, pp. 15–39. Berlin: Springer.

Baker, David 2014. *The Schooled Society: The Educational Transformation of Global Culture*. Stanford: Stanford University Press.

Baker, David, Motoko Akiba, Gerald LeTendre, and Alexander Wiseman. 2001. "Worldwide Shadow Education: Outside-School Learning, Institutional Quality of Schooling, and Cross-National Mathematics Achievement." *Educational Evaluation and Policy Analysis* 23(1): 1–17.

Baker, David, Helmut Kohler, and Manfred Stock. 2007. "Socialist Ideology and the Contraction of Higher Education: Institutional Consequences of State Manpower and Educational Planning in the Former East Germany." *Comparative Education Review* 51(3): 353–377.

Baltagi, Badi 2013. *Econometric Analysis of Panel Data*, (Fifth Edition) edn. West Sussex, UK: John Wiley and Sons.

Barro, Robert and Jong-Wha Lee 2015. *Education Matters: Global Schooling Gains from the 19th to the 21st Century*. New York: Oxford University Press.

Ben-David, Joseph, and Awraham Zloczower. 1962. "Universities and Academic Systems in Modern Societies." *European Journal of Sociology* 3(1): 45–84.

Benavot, Aaron. 1996. "Education and Political Democratization: Cross-National and Longitudinal Findings." *Comparative Education Review* 40(4): 377–403.

Benavot, Aaron and Julia Resnik 2006. "Lessons From the Past: A Comparative Socio-Historical Analysis of Primary and Secondary Education". In *Educating All Children: A Global Agenda*, edited by Cohen, Joel E., Bloom, David E., Malin, Martin B., pp. 123–229. Cambridge, MA: American Academy of Arts and Sciences and the MIT Press.

Bingham, Andrew. 1956. "*Admissions: What Kind of Wheat to Winnow. The Harvard Crimson.*" Accessed 21 June 2019. https://www.thecrimson.com/article/1956/1/6/admissions-what-kind-of-wheat-to/.

Bishop, John. 1997. "The Effect of National Standards and Curriculum-Based Exams on Achievement." *American Economic Review* 87(2): 260–264.

Bol, Thijs and Herman van de Werfhorst 2011. *Measuring Educational Institutional Diversity: External Differentiation, Vocational Orientation, and Standardization*. Amsterdam Working Paper Series: Amsterdam Centre for Inequality Studies.

Bol, Thijs, and Herman van de Werfhorst. 2013. "Educational Systems and the Trade-off between Labor Market Allocation and Equality of Educational Opportunity." *Comparative Educational Review* 57(2): 285–308.

Bol, Thijs, Jacqueline Witschge, Herman van de Werfhorst, and Jaap Dronkers. 2014. "Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 Countries." *Social Forces* 92(4): 1545–1572.

Boli, John and George Thomas (eds) 1999. *Constructing World Culture: International Nongovernmental Organizations Since 1875*. Stanford: Stanford University Press.

Boudon, Raymond 1973. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: John Wiley and Sons.

Bourdieu, Pierre 1996 [1989]. *The State Nobility: Elite Schools in the Field of Power*. Stanford: Stanford University Press.

Bray, Mark. 2017. "Schooling and its Supplements: Changing Global Patterns and Implications for Comparative Education." *Comparative Education Review* 61(3): 469–491.

Breen, Richard and Jan Jonsson 2005. "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility." *Annual Review of Sociology* 31:223–43.

Busemeyer, Marius 2015. *Skills and Inequality: Partisan Politics and the Political Economy of Education Reforms in Western Welfare States*. New York: Cambridge University Press.

Carson, John 2007. *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics*, pp. 1750–940. New Jersey: Princeton University Press.

Chabbott, Colette 2003. *Constructing Education for Development: International Organizations and Education for All*. New York: RoutledgeFalmer.

Clark, Burton and Guy Neave (eds) 1992. *The Encyclopedia of Higher Education: National Systems of Higher Education (Volume 1)*. New York: Pergamom.

Coleman, James 1965. *Education and Political Development*. Princeton: Princeton University Press.

Collins, Randall 1979. *The Credential Society: A Historical Sociology of Education and Stratification*. New York: Academic Press.

Davies, Dan. 2015. "The 'Iron Gate': High-Stakes Assessment at Age 16 in Nepal and England." *Compare: A Journal of Comparative and International Education* 46(4): 582–602.

DiPrete, Thomas, Thijs Bol, Christina Eller, Herman van de Werfhorst. 2017. "School-to-Work Linkages in the United States, Germany, and France." *American Journal of Sociology* 122(6): 1869–1938.

Domina, Thad, Andrew Penner and Emily Penner 2017. "Categorical Inequality: Schools as Sorting Machines." *Annual Review of Sociology* 43:311–30.

Drori, Gili, John Meyer, Francisco Ramirez and Evan Schofer 2003. *Science in the Modern World Polity*. Stanford, CA: Stanford University Press.

Eckstein, Max and Harold Noah 1992. *Examinations: Comparative and International Studies*. New York: Pergamon Press.

Feenstra, Robert, Robert Inklaar, and Marcel Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105(10): 3150–3182.

Foucault, Michel 2009. *Security, Territory, Population*. New York: Picador.

Furuta, Jared. 2017. "Rationalization and Student/School Personhood in U.S. College Admissions: The Rise of Test-Optional Policies, 1987 to 2015." *Sociology of Education* 90(3): 236–254.

Furuta, Jared 2020. *Path Dependence and World Society in National Education Systems: Global Trends in the Use of High-Stakes Examinations, 1960–2010*. Working Paper.

Grodsky, Eric, John Robert Warren and Erika Felts 2008. "Testing and Social Stratification in American Education." *Annual Review of Sociology* 34:385–404.

Hall, Peter and David Soskice (eds.) 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.

Heidenheimer, Arnold 1981. "Education and Social Security Entitlements in Europe and America." In *The Development of Welfare States in Europe and America*, edited by Peter Flora, and Arnold Heidenheimer. Pp. 269–306New York: Oxford University Press.

Helms, Robin Matross 2008. *University Admission Worldwide*. Washington, DC Education Working Paper Series, Number 15: World Bank Group.

Heubert, Jay and Robert Hauser (eds) 1999. *High Stakes Testing for Tracking, Promotion, and Graduation*. Washington, D.C.: National Academy Press.

Heyneman, Stephen, and Angela Ransom. 1990. "Using Examinations and Testing to Improve Educational Quality." *Educational Policy* 4(3): 177–192.

Hironaka, Ann 2014. *Greening the Globe: World Society and Environmental Change*: New York: Cambridge University Press.

Husen, Torsten and Postlethwaite Neville (eds) 1985. *The International Encyclopedia of Education*. Oxford: Pergamon Press.

Jackson, Michelle and Elizabeth Buckner 2016. "Opportunity without Equity: Educational Inequality and Constitutional Protections in Egypt." *Sociological Science* 3:730–56.

Jacob, Brian. 2001. "Getting Tough? The Impact of High School Graduation Exams." *Educational Evaluation and Policy Analysis* 23(2): 99–121.

Kandel, Isaac L. 1930. *History of Secondary Education: A Study in the Development of Liberal Education*. Cambridge, MA: Riverside Press.

Karabel, Jerome 2005. *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Princeton, and Yale*. New York: Houghton Mifflin.

Kellaghan, Thomas and Vincent Greaney 1992. *Using Examinations to Improve Education: A Study in Fourteen African Countries*. Washington, DC World Bank Technical Paper No. 165. Africa Technical Department Series: World Bank.

Kerckhoff, Alan 2001. "Education and Social Stratification Processes in Comparative Perspective." *Sociology of Education* 74:3–18.

Kwang, Tan Yap, Chow Hong Kheng and Christine Goh 2008. *Examinations in Singapore: Change and Continuity*, pp. 1891–2007. Singapore: World Scientific Publishing.

Lauwerys, Joseph and David Scanlon 1969. *The World Yearbook of Education: Examinations*. London: Evans Brothers.

Mahoney, James and Kathleen Thelen 2010. *Explaining Institutional Change: Ambiguity, Agency, and Power*. New York: Cambridge University Press.

Marshall, Monty, Ted Gurr and Keith Jaggers 2014. *Polity IV Project: Political Regime Characteristics and Transitions*, pp. 1800–2013. Vienna, Virginia: Center for Systemic Peace.

Maynes, Mary Jo 1985. *Schooling in Western Europe: A Social History*. Albany: SUNY Press.

Meyer, John. 1977. "The Effects of Education as an Institution." *American Journal of Sociology*, 83(1): 55–77.

Meyer, John, John Boli, George Thomas, and Francisco Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103(1): 144–181.

Meyer, John, Francisco Ramirez, and Yasemin Soysal. 1992. "World Expansion of Mass Education, 1870–1980." *Sociology of Education* 65(2): 128–149.

Mundy, Karen 1999. "Educational Multilateralism in a Changing World Order: UNESCO and the Limits of the Possible." *International Journal of Educational Development* 19:27–52.

Park, Hyunjoon, Claudia Buchmann, Jaesung Choi and Joseph Merry 2016. "Learning beyond the School Walls: Trends and Implications." *Annual Review of Sociology* 42:231–52.

Raban, Sandra 2008. *Examining the World: A History of the University of Cambridge Local Examinations Syndicate*. New York: Cambridge University Press.

Ramirez, Francisco and Richard Rubinson 1979. "Creating Members: The Political Incorporation and Expansion of Public Education". In *National Development and the World System: Educational, Economic, and Political Change, 1950–1970*, edited by Meyer, John, Hannan, Michael, pp. 72–82. Chicago: University of Chicago Press.

Ramirez, Francisco and Marc Ventresca 1992. "Building the Institution of Mass Schooling: Isomorphism in the Modern World". In *The Political Construction of Education: The State, School Expansion, and Economic Change*, edited by Fuller, Bruce, Rubinson, Richard, pp. 47–60. New York: Praeger.

Ramirez, Francisco, John Meyer and Julia Lerch 2016. "World Society and the Globalization of Educational Policy". In *The Handbook of Global Education Policy*, edited by Mundy, Karen, Green, Andy, Lingard, Robert, Verger, Antoni, pp. 43–63. Malden: Wiley-Blackwell.

Ramirez, Francisco, Evan Schofer, and John Meyer. 2018. "International Tests, National Assessments, and Educational Development (1970–2012)." *Comparative Education Review* 62(3): 344–364.

Margaret, Raymond and Eric Hanushek 2003. "High Stakes Research". *Education Next* 3(3): 48–55.

Ringer, Fritz 1979. *Education and Society in Modern Europe*. Bloomington: Indiana University Press.

Rohlen, Thomas 1983. *Japan's High Schools*. Berkeley: University of California Press.

Rubinson, Richard. 1986. "Class Formation, Politics, and Institutions: Schooling in the United States." *American Journal of Sociology* 92(3): 519–548.

Schafer, Mark J. 1999. "International Non-Governmental Organizations and Third World Education in 1990." *Sociology of Education* 72:69–88.

Schofer, Evan, O. Ramirez Francisco and John Meyer. 2005. "The Worldwide Expansion of Higher Education in the Twentieth Century." *American Sociological Review* 70(6): 898–920.

Schofer, Evan, Francisco O. Ramirez and John W. Meyer 2020. "The Societal Consequences of Higher Education." *Sociology of Education*, https://doi.org/10.1177/0038040720942912.

Seth, Michael 2002. *Education Fever: Society, Politics, and the Pursuit of Schooling in South Korea*. Honolulu: University of Hawaii Press.

Solt, Frederic. 2014. "*The Standardized World Income Inequality Dataset*." https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LM4OWF.

Soares, Joseph 1999. *The Decline of Privilege: The Modernization of Oxford University*. Stanford: Stanford University Press.

Stevenson, David Lee, and David Baker. 1992. "Shadow Education and Allocation in Formal Schooling: Transition to University in Japan." *American Journal of Sociology* 97(6): 1639–1657.

Thelen, Kathleen and Sven Steinmo 1992. "Historical Institutionalism in Comparative Politics". In *Structuring Politics: Historical Institutionalism in Comparative Analysis*, edited by Steinmo, Sven, Thelen, Kathleen, Longstreth, Frank. New York: Cambridge University Press.

Thornton, Patricia, and William Ocasio. 1999. "Institutional Logics and the Historical Contingency of Power in Organizations: Executive Succession in the Higher Education Publishing Industry, 1958–1990." *American Journal of Sociology* 105(3): 801–843.

Toenjes, Laurence A., A. Gary Dworkin, Jon Lorence and Antwanette N. Hill 2002. "High-Stakes Testing, Accountability, and Student Achievement in Texas and Houston". In *Bridging the Achievement Gap*, edited by Chubb, John E., Loveless, Tom, pp. 109–30. Washington, DC: Brookings Institution Press.

Turner, Ralph. 1960. "Sponsored and Contest Mobility and the School System." *American Sociological Review*, 25(6): 855–867.

Union of International Associations (UIA). 1960–2010. *Yearbook of International Organizations*. Brussels: Union of International Associations.

United Nations Education, Scientific, and Cultural Organization (UNESCO). 2006. World Data on Education. UNESCO Online Database. UNESCO Institute for Statistics Online Publication.

Warren, John Robert and Krista N. Jenkins 2005. "High School Exit Examinations and High School Dropout in Texas and Florida, 1971–2000." *Sociology of Education* 78:122–43.

Warren, John Robert, and Melanie Edwards. 2005. "High School Exit Examinations and High School Completion: Evidence from the Early 1990's." *Educational Evaluation and Policy Analysis* 27(1): 53–74.

Warren, John Robert, and Rachael Kulick. 2007. "Modeling States' Enactment of High School Exit Examination Policies." *Social Forces* 86(1): 215–229.

Warren, John Robert, Krista Jenkins, and Rachael Kulick. 2006. "High School Exit Examinations and State-Level Completion and GED Rates, 1975-2002." *Educational Evaluation and Policy Analysis* 28(2): 131–152.

Watson, Keith 1982. *Education in the Third World*. New York: Routledge.

Weber, Max 1978. *Economy and Society*. New York: Bedminster Press, Inc.

Weir, Margaret, Ann Orloff and Theda Skocpol 1988. *The Politics of Social Policy in the United States*. Princeton: Princeton University Press.

Windolf, Paul 1997. *Expansion and Structural Change: Higher Education in Germany, the United States, and Japan*, pp. 1870–990. Colorado: Westview Press.

Wooldridge, Jeffrey 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

World Bank 2019. *World Development Indicators (dataset)*: World Bank Publication.

World Bank. 2020. "*How Can Gross School Enrollment Ratios Be over 100 Percent?*" Accessed February 18, 2020. https://datahelpdesk.worldbank.org/knowledgebase/articles/114955-how-can-gross-school-enrollment-ratios-be-over-100.

World Education Forum 2000. "The Dakar Framework for Action". In *Education for All: Meeting Our Collective Commitments. Adopted by the World Education Forum (Dakar, Senegal)*. Paris: UNESCO.

Zapp, Mike and Francisco O. Ramirez 2018. *Beyond Internationalization and Isomorphism: The Construction of a Global Higher Education Regime*. Working Paper.