

Research statement — Evan Riehl, Cornell University

As a labor economist focusing on the economics of education, my research centers on two main questions: 1) how and why do the labor market returns to attending selective colleges vary across students and 2) how does the use of standardized exams in education affect the incentives and outcomes of students and schools. My work uses reduced-form methods to tease out causal impacts from real world, often large scale policies. This methodological approach allows me to examine mechanisms that may not arise in experimental or smaller scale studies while also relying on weaker assumptions than those commonly employed in structural work. Many of my papers utilize data from Colombia and Brazil which, like the United States, employ a decentralized and privatized approach to higher education, but which often implement more ambitious public policy. Examining this data allows me to investigate the actual labor market impacts of various major reforms to the higher education system.

In particular, my research informs current debates about the use of standardized exams and affirmative action in college admissions. Concerns about bias in standardized exams have led a growing number of U.S. colleges to switch to test-optional or test-blind admissions. Affirmative action is contentious because it gives applicants from disadvantaged backgrounds admission priority over other applicants with better academic credentials. These debates have become greatly intensified with the U.S. Supreme Court's June 2023 ban on race-based affirmative action as selective universities will be forced to change admission policies. It is therefore crucial to provide compelling evidence about the actual economic impact of these policies in order to facilitate sound policy decisions. Most existing work focuses on the outcomes of students who are marginally admitted to selective colleges, but this may fail to capture the full impacts of significant reforms to admission policy. My work fills this evidence gap by examining both the direct labor market impacts of admission policies on students who gain or lose access to selective colleges as well as the spillover effects of these policies on other students in the education system.

Several of my papers show that admission policies that give preference to disadvantaged students can have negative consequences because selective universities derive part of their value from having high-achieving student bodies. In Riehl (2023, forthcoming at *JOLE*), I study a reform of the Colombian national college admission test that significantly reduced the exam's predictive power for student success. I show that this reform allowed lower-income students to attend more selective colleges, but it reduced systemwide graduation rates and earnings for *both* high- and lower-income students. In Machado, Reyes, and Riehl (2023, forthcoming at *JOLE*), we examine an affirmative action policy at an elite Brazilian university that reserved 45 percent of admission slots for Black and low-income students. We find that the policy had modest benefits for affirmative action students, but it had large negative spillover effects on the labor market outcomes of the university's other students. These papers show that policies that decrease the concentration of high-achieving students in selective universities can inadvertently reduce the productivity of these schools and the higher education system as a whole.

To shed light on the causes of these findings, my research has also explored some of the mechanisms involved in the returns to college selectivity. In MacLeod, Riehl, Saavedra, and

Urquiola (2017, *AEJ:Applied*), we show that college reputation—defined as the *expected* skill of a college’s graduates—has a causal impact on individual earnings outcomes. In de Roux and Riehl (2022b, *JPubE*), we show that college students are more likely to fail courses and drop out when they are relatively less prepared than their classmates. These papers can explain some of my findings in Riehl (2023) and Machado, Reyes, and Riehl (2023), and they illustrate why selective universities have historically given admission preference to high-achieving students.

A key policy implication from my research is that it is difficult to reduce earnings inequality between advantaged and disadvantaged students through college admissions alone. Since my work finds that academic preparation is important for students to succeed at selective universities, it is crucial for policy to address the sources of achievement gaps that arise at younger ages. This motivated my work showing that community exposure to mass incarceration is one factor that contributes to the achievement gaps of U.S. children (Gupta, Hansman, and Riehl, 2023, Working Paper). Yet an important exception to this conclusion comes from my work in Ng and Riehl (2023, accepted at *AEJ:Policy*), which shows that the average earnings returns to enrolling in selective Science, Technology, Engineering, and Math (STEM) programs are larger for less-prepared students than for their more-prepared classmates, despite the fact that less-prepared students are more likely to drop out of these programs. Thus policies that support disadvantaged students in pursuing STEM degrees may be an effective means of reducing earnings inequality.

A complementary line of my work asks how the design and use of standardized exams affects incentives and information in the education system. In Riehl and Welch (2023, *JPAM*), we show that the incentives for educators to engage in test prep tend to be stronger in math than in English/language arts (ELA) due to differences in the design of standardized exams. This can partly explain a common finding that has puzzled education researchers—namely, that educational inputs tend to affect math scores more than ELA scores in settings where accountability matters. In Reyes, Riehl, and Xu (2023, Working Paper), we show that test prep for high-stakes college admission exams makes the exam scores *more* informative for students’ outcomes, although it also widens test score gaps between high- and lower-income students. This is consistent with the common criticism that high-stakes exams give wealthy students a leg up in admissions, but it contradicts the claim that test prep causes exam scores to be biased as a measure of academic potential.

Below I describe all of my research and its contributions in more detail.

1 Returns to selective colleges

A major line of my research examines the earnings returns to attending a selective college. Admission slots at selective universities are a scarce resource, and there is a perennial debate over who can and should benefit from this resource. Selective universities typically spend much more money per student than other institutions (Hoxby, 2009), and students from disadvantaged backgrounds may particularly benefit from this spending. But disadvantaged students also tend to be less prepared academically, so they may struggle at schools with more advanced curricula

(Arcidiacono, Aucejo, and Hotz, 2016). Further, the value of attending a selective college may depend on the composition of its students. Universities value diversity as an important component of their students' educational experiences, and yet they also perceive that high achieving student bodies will help them bring the best faculty and employers to campus.

My work sheds light on this debate by addressing two main questions: 1) how do the returns to college selectivity vary among students with different levels of academic preparation? and 2) what are the mechanisms for these returns?

The role of academic preparation

In “Do less informative college admission exams reduce earnings inequality? Evidence from Colombia” (Riehl, 2023, forthcoming at *JOLE*), I examine a topical issue in the debate over selective university admissions: the role of standardized admission exams. Historically and across the globe, standardized exams have been the primary mechanism that universities have used to identify and admit high-achieving students. More recently, these exams have come under fire as instruments that measure wealth and privilege more than academic potential. These criticisms have led many selective universities in the United States to switch to test-optional admissions or to forgo using exam scores altogether.¹

My paper asks how reducing the influence of exam scores in admissions affects labor market outcomes for both advantaged and disadvantaged students. I examine the Colombian national college admission test, known as the ICFES, which underwent a major overhaul in the year 2000 with the goal of reducing socioeconomic bias in scores. The overhaul included several changes that reduced the exam's information content, such as decreases in the number of multiple-choice options and in the total number of questions. I show that the ICFES reform dramatically reduced test score gaps between high- and low-income students, but it did so primarily by making the exam a worse predictor of college GPAs and graduation rates.

Using regional variation in universities' use of the ICFES exam in admissions, I show that the exam reform allowed low-income students to attend more selective colleges, while higher-income students were displaced to less selective schools. The magnitude of this change in the sorting of students to college was substantial. In areas where the ICFES mattered in admissions, the reform enabled roughly one in 20 low-income students to attend the next most selective college in their region, while roughly one in 20 higher-income students moved down a selectivity tier.

My main finding is that the ICFES reform reduced graduation rates and earnings for *both* high- and low-income students. In affected regions, post-college earnings fell by 1–2 percent for both high- and low-income students relative to earlier cohorts. These negative effects persisted up through 16 years after taking the ICFES exam. Graduation rates also declined for both groups, suggesting that the reform caused students to attend colleges where they were less likely to succeed academically.

My findings show that reforms that lessen the influence of exam scores in admissions can

¹For example, the University of Chicago switched to test-optional admissions in 2018, and, in 2021, the University of California announced that it would no longer consider SAT or ACT scores at all.

inadvertently harm disadvantaged students by reducing the quality of student/college matches. My paper differs from studies that find that disadvantaged students benefit from attending selective colleges using regression discontinuity designs (e.g., Zimmerman, 2014; Smith, Goodman, and Hurwitz, 2020) or randomized controlled trials (Angrist, Autor, and Pallais, 2022). By the nature of their empirical strategies, these papers estimate returns for students whom the universities deemed qualified for admission. By contrast, the earnings effects of the ICFES reform depend on students who were *not* qualified under the old exam, and who may have lower returns. In this sense, my paper is more similar to work that estimates students' counterfactual returns at different colleges (e.g., Arcidiacono, Aucejo, and Hotz, 2016; Dillon and Smith, 2020), but my results rely on weaker assumptions.

Further, my results show that policies that reduce the influence of exam scores in admissions can reduce the overall productivity of the education system. In the United States, these policies are currently being debated and implemented at each of the K–12, undergraduate, and graduate schooling levels. For example, the Educational Testing Service recently announced that the GRE—a widely used U.S. graduate admission test—will now take only two hours to complete instead of four to lessen the burden on test takers.² Like the Colombian ICFES reform, this change to the GRE will reduce the exam's information content, and it may therefore have negative consequences for the quality of matches between students and graduate programs.

My work in Riehl (2023) was unable to separately identify direct effects on students who changed colleges as a result of the reform from spillover effects on other students in the higher education system. My co-authors and I take on this challenge in **“The direct and spillover effects of large-scale affirmative action at an elite Brazilian university” (Machado, Reyes, and Riehl, 2023, forthcoming at *JOLE*)**. This paper speaks to the ongoing and high-profile debate on the desirability of affirmative action in college admissions. The context for our paper is an elite public university in Brazil called Rio de Janeiro State University, or UERJ for short. UERJ was one of the first universities in Brazil to adopt affirmative action in admissions, and it did so on a large scale. Beginning in 2004, UERJ implemented both race- and class-based quotas that governed admissions for 45 percent of the seats in each major.

Our paper is novel in identifying both the direct effects of affirmative action on its intended beneficiaries and the policy's spillover effects on other students. To identify direct effects on students admitted through affirmative action, we follow other work on college selectivity in using a regression discontinuity (RD) design that compares applicants above and below admission score cutoffs (e.g., Hoekstra, 2009). Uniquely, we identify spillover effects on other students by exploiting the fact that the reserved quotas often did not fill up in some of UERJ's less-prestigious majors. This meant that affirmative action students made up 45 percent of the incoming class in UERJ's most selective majors, whereas affirmative action students comprised only 10–20 percent of the class in other majors. We implement a difference-in-differences design that estimates the impacts of greater exposure to affirmative action in a sample of *top enrollees*, which we define as students whose entrance exam scores were high enough to gain admission

²See: “Shorter GRE Test Coming September 2023,” Educational Testing Service.

regardless of whether affirmative action existed in their cohort.

We find that UERJ’s policy had modest benefits for students admitted through affirmative action, but it had negative spillover effects on the earnings outcomes of UERJ’s top enrollees. Enrolling in UERJ led to a 14 percent increase in affirmative action students’ early-career wages, although these wage gains faded as their careers progressed. Conversely, the average wages of top enrollees decreased by 14 percent in majors with high affirmative action take-up relative to those with lower take-up, and this effect persisted up through 13 years after enrollment. Using employer-employee data and records from a college graduation exam, we show that the negative spillover effects on top enrollees were partly driven by both networking and learning mechanisms. Specifically, the adoption of affirmative action reduced the likelihood that top enrollees obtained jobs at high-paying firms affiliated with UERJ peers and alumni, and it reduced their performance on the graduation exam.

There is limited research on the earnings impacts of affirmative action, but our findings tell a more cautionary tale about the efficacy of these policies than in two recent *QJE* papers. Chetty, Friedman, Saez, Turner, and Yagan (2020) estimate returns to attending many U.S. colleges, and they use these estimates to argue that large-scale income-based affirmative action could significantly increase intergenerational mobility. Bleemer (2022) shows that affirmative action in the University of California (UC) system raised minority students’ earnings. Our estimates of the earnings benefits of affirmative action for disadvantaged students are substantially smaller than in both of these papers. In addition, neither paper examines spillover effects of affirmative action on other students, which we find to be negative and substantial in magnitude.³

Our results show that selective universities derive part of their value from admitting high-achieving and high-income students, and thus large-scale affirmative action can reduce these benefits for all students, including those from disadvantaged backgrounds. There are many ways in which the value of attending a top university can depend on its student composition, including the level at which faculty can teach courses, the employers that are attracted to campus, and the value of networking with peers and alumni. Our results illustrate why universities have historically given preference to high-achieving and high-income students, as well as the potential negative effects they must be aware of and mitigate in the effort to diversify their classes.

Two of my other papers are similar in finding that academic preparation is an important determinant of college success. In “**Disrupted academic careers: The returns to time off after high school**” (de Roux and Riehl, 2022a, *JDE*), we estimate the returns to a college education by exploiting an academic calendar transition in Colombia that caused some high school graduates to forgo enrolling in college. Students with high levels of academic preparation experienced decreases in earnings from this policy, which suggests that they would have had high returns to college enrollment. Conversely, we find that many of the less-prepared students who were affected would have dropped out of college if they had enrolled, and thus forgoing college did not significantly impact their earnings outcomes.

³Chetty, Friedman, Saez, Turner, and Yagan (2020) write: “[W]e also assume that [our] estimated causal effects do not change under our counterfactual student reallocations, in particular ignoring potential changes in value-added that may arise from having a different group of students (peer effects)” (p. 1626).

In “**Do college students benefit from placement into higher-achieving classes?**” (de Roux and Riehl, 2022b, *JPubE*), we examine a unique admission policy at a Colombian flagship university called Univalle (Universidad del Valle) that placed students into either high- or lower-achieving sections of the same classes. We find that students who were marginally placed into the high-achieving sections were more likely to fail the courses and drop out of the program than students who placed into the lower-achieving sections. Contrary to work at the K–12 level (Card and Giuliano, 2016), this shows that college students can benefit when they are *relatively* more prepared academically than many of their classmates.

Yet this pattern of findings flips in my final paper in this topic area, “**The returns to STEM programs for less-prepared students**” (Ng and Riehl, 2023, *accepted at AEJ:Policy*). This paper asks how academic preparation influences the returns to enrolling in STEM programs at selective universities. This focus is motivated by policies that aim to attract more students into STEM programs. These policies affect individuals on the margin of STEM enrollment, who tend to be less prepared for STEM coursework than inframarginal enrollees. Existing work finds that there are large average returns to a *completed* STEM degree (e.g., Kirkebøen, Leuven, and Mogstad, 2016), but other work shows that less-prepared students are more likely to drop out of STEM programs (e.g., Stinebrickner and Stinebrickner, 2014). Our innovation in this paper is to jointly examine the relationship between students’ completion rates and earnings returns—leading to a more complete analysis of the impacts of expanding STEM education to less-prepared student populations.

We use data from Univalle—the Colombia flagship university—and two novel empirical strategies to estimate how the returns to enrolling in selective STEM programs vary with students’ academic preparation. Our first strategy takes advantage of data on a large number of pre-college test scores, which allows us to estimate heterogeneity in the RD effects of crossing Univalle admission thresholds based on an applicant’s predicted likelihood of completing a STEM degree. Our second strategy exploits expansions in Univalle’s STEM quotas in some majors and cohorts, which allowed some less-prepared applicants to gain admission.

Our main finding is that less-prepared students had significantly *higher* earnings returns to enrolling in Univalle’s STEM programs than more-prepared students. Across our two empirical strategies, we find that less-prepared students experienced 30–40 percent earnings gains from enrolling in Univalle’s STEM program compared to their counterfactual college options. By contrast, the returns for more-prepared students were close to zero. The higher returns for less-prepared students occurred despite the fact that they were 9–18 percentage points less likely to complete the Univalle STEM program than more-prepared enrollees.

I was surprised by these findings given my other work in this area, but this paper sheds light on the contexts in which one would expect disadvantaged students to benefit from admission to selective colleges. A key mechanism for the findings in Ng and Riehl (2023) is that, relative to more-prepared students, less-prepared students were less likely to enroll in STEM programs at other universities if they were rejected from Univalle. This shows that admission or scholarship policies that allow less-prepared students to enroll in selective STEM programs may be more

helpful in reducing earnings inequality than similar policies for other selective degree programs like business or law (e.g., Zimmerman, 2019). Our findings may also be due in part to an undersupply of high-quality undergraduate STEM programs in Colombia, which argues for increased funding to support the creation of new STEM programs.

On a more cautionary note, the estimates in Ng and Riehl (2023) reflect the returns for marginally-admitted students given the status quo composition of students in Univalle’s STEM programs, so our findings may not generalize to policies that significantly alter peer composition. This can explain why the findings in Ng and Riehl (2023) are qualitatively different than my findings in Riehl (2023) and Machado, Reyes, and Riehl (2023). Taken together, these papers highlight a key takeaway from my work: researchers should be cautious about using estimates for marginal college admits to make predictions on larger-scale reforms to higher education.

Mechanisms

A related line of my research examines mechanisms for the returns to college selectivity. It is often said that part of the benefit of attending a selective university comes from the school’s reputation, but there was little compelling evidence on this channel prior to my work in **“The big sort: College reputation and labor market outcomes” (MacLeod, Riehl, Saavedra, and Urquiola, 2017, *AEJ:Applied*)**. Economists have a precise definition of reputation; namely, a firm’s reputation is the expected quality of its product (Shapiro, 1983). In higher education, one can view colleges as “firms” and employers as the “buyers” of their product. Thus a college’s reputation can be defined as the expected skill of its graduates.

Our paper adopts a tractable definition of a college’s reputation—the mean admission score of its graduates—and shows that reputation has a causal effect on individuals’ post-college earnings. Using nationwide data from Colombia, we show that earnings and college reputation are positively correlated after controlling for an individual’s own admission score. This correlation may arise because of unobserved sorting into colleges, because high-reputation colleges provide more skill, or because college identity signals graduates’ ability. To differentiate between these mechanisms, we exploit the staggered introduction of a national college graduation exam, which provided employers with a new signal of individual skill. We develop a competitive labor market model (e.g., Jovanovic, 1979; Farber and Gibbons, 1996; Altonji and Pierret, 2001) and show that the graduation exam should reduce the correlation of earnings with college reputation if reputation serves as a signal of skill. We find exactly this in our empirical analysis.

Our analysis sheds light on mechanisms for the findings in many papers on college selectivity, including my own work. If reputation is important, then a student who is marginally admitted to a college program with many high-achieving students may experience better labor market outcomes. This may be partly behind my findings in Ng and Riehl (2023), as well as the findings in many papers that use RD designs to estimate returns to college selectivity. At the same time, admission policies that reduce the average skill of a college’s student body can reduce the earnings of inframarginal students, as in, for example, Machado, Reyes, and Riehl (2023).

Machado, Reyes, and Riehl (2023) also provides evidence that networking is an important

mechanism for the value of attending an elite university. We use employer-employee data to show that admission to UERJ significantly increased the likelihood that individuals worked at firms with other UERJ alumni, especially those from their own enrollment cohort. This finding mirrors results in Zimmerman (2019) and Michelman, Price, and Zimmerman (2022), but unlike in these papers, we find that disadvantaged students also benefit from networking at elite universities. This difference can be explained by the fact that our data measures early-career employment outcomes in a broad set of firms, which allows us to capture the types of jobs that disadvantaged students can obtain through networking.

My book chapter, “**Learning and earning: An approximation to college value added in two dimensions**” (Riehl, Saavedra, and Urquiola, 2019, Chapter in *Productivity in Higher Education*, edited by C. Hoxby and K. Stange), also highlights the different ways in which colleges impact student outcomes. In this work, we show that measures of college value added based on earnings are only modestly correlated with value added measures based on learning (as defined by performance on a college graduation exam). Our findings suggest that measures of learning value added may do a better job in isolating a college’s contributions to students’ human capital, while measures of earnings value added are more likely to reflect non-human capital mechanisms like reputation or networking.

2 Incentives and information from the use of standardized tests

Building on my work in Riehl (2023) and MacLeod, Riehl, Saavedra, and Urquiola (2017), a related line of my research asks how the use of standardized tests affects incentives and information in the education system. In K–12 research, economists frequently use standardized test scores as an outcome variable to measure the impacts of educational inputs or policy interventions. Yet a test score is a summary statistic that reflects a student’s average performance across a range of exam questions. These questions can vary widely in the skills that they measure and, therefore, in their informativeness for other outcomes. As a researcher who focuses primarily on higher education, I often wonder whether the test score impacts that I see in one paper are equivalent to those in another paper in terms of their implications for longer-term outcomes. This motivated me to get “under the hood” of standardized tests and view them as an *input* to behavior and economic outcomes.

This alternative approach to standardized exams is exemplified in my paper “**Accountability, test prep incentives, and the design of math and English exams**” (Riehl and Welch, 2023, *JPAM*). I was inspired to write this paper by a pattern that I observed in papers and conference presentations on K–12 topics: Researchers often find that educational inputs have a bigger impact on math scores than on English/language arts (ELA) scores in settings where educators are held accountable for student performance. In our paper, we include a review of three literatures—accountability policies (e.g., Dee and Jacob, 2011); charter schools (e.g., Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak, 2011); and teacher value added (e.g., Chetty, Friedman, and Rockoff, 2014a)—that shows just how common this pattern is. A common hypothesis for this pattern is that most math learning takes place at school,

while students primarily learn English and reading at home. But Jackson, Rockoff, and Staiger (2014) acknowledge: “There is no clear explanation for this fact.”

Our contribution is to show that this pattern can be partly explained by differences in the design of math and ELA exams, which gives educators a stronger incentive for test prep in math than in ELA. We collected data from the technical reports of grades 3 through 8 standardized exams in six states that are the setting for most research on education accountability in the United States. Using this data, we show that math exams, relative to ELA exams, tend to measure ability more precisely for students who are near the margin of achieving proficiency, which is the key metric in U.S. state accountability systems.⁴ We develop a simple theoretical model to show that these differences in exam design created stronger incentives for test prep in math than in ELA. Intuitively, teachers have an incentive to target test prep to marginally-proficient students (Neal and Schanzenbach, 2010), and math exams are more likely to reward teachers for this targeted test prep because they measure ability more precisely for these students. Our empirical simulations show that this difference in incentives is quantitatively meaningful: the effect of targeted instruction on the proficiency rate in an average classroom is 22 percent larger in math than in ELA.

Our paper provides a new explanation for why test score impacts often vary across subjects and/or exams. We show that the math/ELA ratio of point estimates in the papers in our literature review is strongly correlated with the math/ELA ratio of test prep incentives, which suggests that test prep plays a role in these findings. For future researchers, we provide guidance for gauging how test prep incentives vary across exams.

Given the nature of the data in Riehl and Welch (2023), a key question that we could not answer is whether test prep is useful beyond the context of the exam. I take on this question in a new working paper, “**Do high stakes muddle the information from standardized tests? Evidence from Brazil’s ENEM exam**” (Reyes, Riehl, and Xu, 2023). Our paper empirically tests two common criticisms of high-stakes college admission exams. First, critics argue that high-stakes exams help wealthy student “game the system” because they have greater access to test prep services. Second, critics contend that this gaming creates “bias” in scores in the sense that it is not informative about individuals’ academic potential.

We exploit a unique natural experiment in Brazil to empirically evaluate these two criticisms. From 2009–2017, the Brazilian government rolled out a standardized admission exam called the ENEM for its system of highly selective federal universities. The ENEM replaced these schools’ own admission exams as the sole admission criterion. Yet the ENEM also functions as a high school accountability test, so many high school seniors took the ENEM regardless of its use in college admissions. This allows us to define a sample of students who were likely to take ENEM regardless of its stakes, and then ask how the distribution and informativeness of scores changed when top universities adopted the exam in admissions.

We find that high-stakes exams do give wealthy students a leg up in admissions, but that

⁴This difference in precision is explained by several structural features of the tests. For example, math exams tend to have more questions overall, and the difficulty of these questions tends to be better aligned with the ability of marginally-proficient students.

higher-stakes exams are *more* informative for academic potential than lower-stakes exams. Specifically, test score gaps between high- and lower-income students widened when federal universities adopted the ENEM in admissions. Yet we also find that the correlation between ENEM scores and students’ college outcomes increased on the higher-stakes exam.

Our paper is novel in showing that higher exam stakes improve the informativeness of scores. It is often argued that high-stakes testing distorts learning toward test prep skills that are not useful beyond the exam (e.g., Jacob, 2005). Our results show that, at least in the context of admission exams, higher stakes can be beneficial in identifying individuals who are likely to succeed academically. Students who earn better scores on high-stakes exams may have characteristics that help them in other academic contexts, such as the ability to focus or the capacity to learn new material. Increasingly, U.S. colleges are reducing their reliance on high-stakes exams in favor of lower-stakes admission signals like high school grades. Our findings show that this change will help colleges to diversify their student bodies. But if colleges also want to admit students who can succeed in their programs, they must find other admission signals that replace the information content of high-stakes exams.

3 Mass incarceration and children’s academic achievement

Many of my papers find that academic preparation is important for students to succeed in selective colleges, which in turn can limit the effectiveness of policies that seek to reduce inequality through college admissions. This had led me to investigate factors that cause academic achievement to diverge between advantaged and disadvantaged students at younger ages.

In “**Community impacts of mass incarceration**” (Gupta, Hansman, and Riehl, 2023, Working Paper), we ask how the United States’ legacy of high incarceration rates has affected the achievement of American children. Existing research in economics has found that the incarceration of a parent who is on the margin of receiving this sentence can actually have *positive* impacts on their children’s later-life outcomes (e.g., Arteaga, 2021; Norris, Pecenco, and Weaver, 2021). Our paper takes a different approach by asking how mass incarceration affects *all* children in the community. This approach is consistent with work in other fields that highlights how incarceration can negatively impact community networks, norms, and labor markets (e.g., Clear, 2008). Further, as we show in the paper, children are much more likely to be indirectly exposed to incarceration through classmates or neighbors than to be directly exposed through the incarceration of a household member.

We develop a novel empirical strategy that allows us to estimate the causal impacts of community-level incarceration rates on the academic outcomes of North Carolina children. Our approach exploits the entry and exit of judges who differ in stringency (their tendency to incarcerate criminal defendants) in each county court. We show that the turnover of more- or less-stringent judges leads to significant changes in aggregate incarceration rates at the county level but is unrelated to trends in criminal activity and family demographics.⁵ Our analysis is

⁵Our judge turnover design was inspired by an empirical strategy that I knew from the education literature. Chetty, Friedman, and Rockoff (2014b) exploit the arrival and departure of teachers who vary in value added

made possible by a new dataset that links criminal justice and education records for the entire state of North Carolina that we created by matching defendant and student addresses.

Our main finding is that increases in local incarceration rates lead to meaningful reductions in the average achievement of children in those communities. A 20 percent increase in the number of incarcerations in a county reduces the *average* child’s performance on math and English exams by 2–3 percent of a standard deviation. These effects are persistent, growing in magnitude over the course of at least three years. Crucially, they are sizable even when considering students whose households have no direct exposure to the criminal justice system. This suggests that the aggregate consequences of incarceration must be driven, at least in part, by spillovers onto a broad set of children in the community. We conduct auxiliary analyses to provide evidence on one of the many potential mechanisms for these spillovers. Consistent with research on behavioral disruptions in the classroom (e.g., Lazear, 2001; Carrell and Hoekstra, 2010), we find that the incarceration of a household member increases misbehavior rates for directly-impacted children, which reduces the academic performance of their classmates.

Our results show that mass incarceration has contributed to achievement gaps between U.S. children who grew up in more- and less-affluent areas. In particular, community exposure to incarceration may be one factor that contributes to low mobility rates for Black and lower-income children (Chetty, Hendren, Jones, and Porter, 2020). Although a full analysis of the costs and benefits of incarceration policy is beyond the scope of our paper, our findings suggest that recent policy initiatives to decriminalize some offenses and use alternatives to incarceration may help to reduce inequality in children’s academic achievement.

4 Ongoing and future work

Since my existing work has found that it is essential to address inequality in human capital development at younger ages, my ongoing research is focusing on the role of colleges in training teachers. One ongoing project uses data from the Texas Education Research Center to examine the efficacy of for-profit teacher training programs and teacher certification exams. Importantly, this project asks how training and certification requirements affect both the quantity and quality of teachers as measured by their impacts on student achievement. (For details, see the research proposal included with this packet.) In another project, I am putting together nationwide data from Brazil to ask how affirmative action in undergraduate teaching degree programs affected the sorting of teachers to schools and the achievement of students in those schools.

Lastly, in work in progress, I am extending Machado, Reyes, and Riehl (2023) to ask how UERJ’s affirmative action policy affected wages, promotions, and turnover rates in industries and occupations that typically hired graduates from Rio de Janeiro universities. There is little research on how policies that increase the representation of disadvantaged students in selective colleges affect the quality of matches between students and jobs, which is essential for a full understanding of the desirability of affirmative action. My project aims to fill this gap.

to estimate impacts on school-level test scores. Analogously, we exploit the arrival and departure of judges who vary in stringency to estimate impacts on county-level outcomes.

Authored research

- de Roux, N. and E. Riehl (2022a). Disrupted academic careers: The returns to time off after high school. *Journal of Development Economics* 156, 102824.
- de Roux, N. and E. Riehl (2022b). Do college students benefit from placement into higher-achieving classes? *Journal of Public Economics* 210, 104669.
- Gupta, A., C. Hansman, and E. Riehl (2023). Community impacts of mass incarceration. Working Paper.
- Machado, C., G. Reyes, and E. Riehl (2023). The direct and spillover effects of large-scale affirmative action at an elite Brazilian university. *Journal of Labor Economics*. Forthcoming.
- MacLeod, W. B., E. Riehl, J. E. Saavedra, and M. Urquiola (2017). The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics* 9(3), 223–261.
- Ng, K. and E. Riehl (2023). The returns to STEM programs for less-prepared students. *American Economic Journal: Economic Policy*. Accepted.
- Reyes, G., E. Riehl, and R. Xu (2023). Do high stakes muddle the information from standardized tests? evidence from Brazil’s ENEM exam. Working Paper.
- Riehl, E. (2023). Do less informative college admission exams reduce earnings inequality? evidence from Colombia. *Journal of Labor Economics*. Forthcoming.
- Riehl, E., J. E. Saavedra, and M. Urquiola (2019). Learning and earning: An approximation to college value added in two dimensions. In C. Hoxby and K. Stange (Eds.), *Productivity in Higher Education*, Chapter 4, pp. 105–132. The University of Chicago Press.
- Riehl, E. and M. Welch (2023). Accountability, test prep incentives, and the design of math and English exams. *Journal of Policy Analysis and Management* 42(1), 60–96.

Other references

- Abdulkadiroğlu, A., J. D. Angrist, S. M. Dynarski, T. J. Kane, and P. A. Pathak (2011). Accountability and flexibility in public schools: Evidence from Boston’s charters and pilots. *The Quarterly Journal of Economics* 126(2), 699–748.
- Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics* 116(1), 313–350.
- Angrist, J., D. Autor, and A. Pallais (2022). Marginal effects of merit aid for low-income students. *The Quarterly Journal of Economics* 137(2), 1039–1090.

- Arcidiacono, P., E. M. Aucejo, and V. J. Hotz (2016). University differences in the graduation of minorities in STEM fields: Evidence from California. *American Economic Review* 106(3), 525–562.
- Arteaga, C. (2021, 10). Parental Incarceration and Children’s Educational Attainment. *The Review of Economics and Statistics*, 1–45.
- Bleemer, Z. (2022). Affirmative action, mismatch, and economic mobility after California’s Proposition 209. *The Quarterly Journal of Economics* 137(1), 115–160.
- Card, D. and L. Giuliano (2016). Can tracking raise the test scores of high-ability minority students? *American Economic Review* 106(10), 2783–2816.
- Carrell, S. E. and M. L. Hoekstra (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone’s kids. *American Economic Journal: Applied Economics* 2(1), 211–228.
- Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *The American Economic Review* 104(9), 2593–2632.
- Chetty, R., J. N. Friedman, and J. E. Rockoff (2014b). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review* 104(9), 2633–79.
- Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics* 135(3), 1567–1633.
- Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2020). Race and economic opportunity in the united states: An intergenerational perspective. *The Quarterly Journal of Economics* 135(2), 711–783.
- Clear, T. R. (2008). The effects of high imprisonment rates on communities. *Crime and justice* 37(1), 97–132.
- Dee, T. S. and B. Jacob (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management* 30(3), 418–446.
- Dillon, E. W. and J. A. Smith (2020). The consequences of academic match between students and colleges. *Journal of Human Resources* 55(3), 767–808.
- Farber, H. S. and R. Gibbons (1996, November). Learning and wage dynamics. *Quarterly Journal of Economics* 111(4), 1007–47.
- Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics* 91(4), 717–724.

- Hoxby, C. M. (2009). The changing selectivity of American colleges. *The Journal of Economic Perspectives* 23(4), 95–118.
- Jackson, C. K., J. E. Rockoff, and D. O. Staiger (2014). Teacher effects and teacher-related policies. *Annual Review of Economics* 6(1), 801–825.
- Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics* 89(5-6), 761–796.
- Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy* 87, 972–90.
- Kirkebøen, L., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics* 131(3), 1057–1111.
- Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics* 116(3), 777–803.
- Michelman, V., J. Price, and S. D. Zimmerman (2022). Old boys’ clubs and upward mobility among the educational elite. *The Quarterly Journal of Economics* 137(2), 845–909.
- Neal, D. and D. W. Schanzenbach (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics* 92(2), 263–283.
- Norris, S., M. Pecenco, and J. Weaver (2021, September). The effects of parental and sibling incarceration: Evidence from ohio. *American Economic Review* 111(9), 2926–63.
- Shapiro, C. (1983). Premiums for high quality products as returns to reputations. *Quarterly Journal of Economics* 98(4), 659–679.
- Smith, J., J. Goodman, and M. Hurwitz (2020). The economic impact of access to public four-year colleges. National Bureau of Economic Research.
- Stinebrickner, R. and T. R. Stinebrickner (2014). A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies* 81(1), 426–472.
- Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics* 32(4), 711–754.
- Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review* 109(1), 1–47.