

Who Should Count as an Expert in K-12 Students' Needs? Parents versus Predictive Models¹

Rebecca A. Johnson ¹, Simone Zhang ², Katherine Christie ³

¹Assistant Professor, Georgetown McCourt School of Public Policy

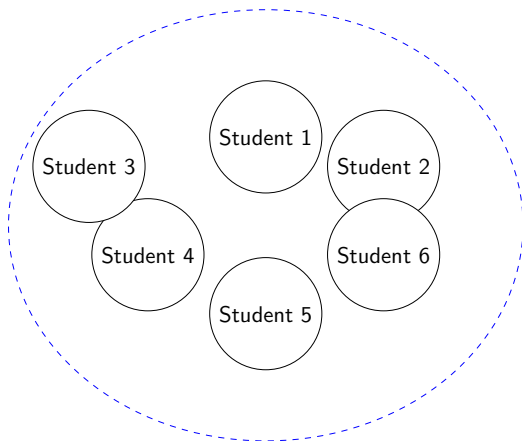
²Assistant Professor, University of Notre Dame Sociology

³BA, Dartmouth College

McCourt Colloquium; 10.25.22

¹Thanks to the following students for excellent research assistance—Shengrou (Rosy) Zhong, Liz Moison, and Morgan Welch—and to Leah Jones for a related collaboration. Funding comes from the Dartmouth Neukom Institute for Computational Science, the NSF TESS Young Investigators Special Competition, and the Spencer/NAEd Postdoctoral Fellowship. A pre-analysis plan that contains the detailed survey design is here: <https://osf.io/fjb56/>.

Focus: predictive algorithms to rank student “need for help”



Contrasts: algorithms to judge **risk to society**
or **value to organizations**

Counterfactual perspective on fairness

RESEARCH-ARTICLE OPEN ACCESS



What is the Bureaucratic Counterfactual? Categorical versus Algorithmic Prioritization in U.S. Social Policy

Authors:  [Rebecca Ann Johnson](#),  [Simone Zhang](#) [Authors Info & Claims](#)

FACCT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency • June 2022 • Pages 1671–1682 • <https://doi.org/10.1145/3531146.3533223>

Expertise about need

- ▶ Past research examines tension between a limited set of counterfactuals
 - ▶ **Human judgment with full discretion**
 - ▶ **Algorithmic/quantitative predictions**
- ▶ Two ways we complicate
 - ▶ **Human judgment:** whose judgment takes precedence?
 - ▶ **Quantitative predictions:** contrast between manually-chosen, coarse categories (e.g., veteran; $\text{income} \leq \text{threshold}$) and inductively generated predictions of some target outcome

Case study: K-12 districts judging need for “high-dosage/high impact tutors” in wake of COVID-19

STUDENT ACHIEVEMENT EXPLAINER

High-Dosage Tutoring Is Effective, But Expensive. Ideas for Making It Work

Existing research: which SEAs/LEAs are using ESSER funds for tutoring² or impact of tutoring on student outcomes³

²Future Ed; Edunomics Lab

³National Student Support Accelerator

Case study: K-12 districts judging need “high-dosage/high impact tutors” in wake of COVID-19

Ways of targeting in-person tutoring resources:

- ▶ “Students, identified through data, who have been most impacted by COVID-19”⁴
- ▶ “Students who were English language learners, students with disabilities, students with a history of chronic absenteeism, and students who were struggling in coursework before the pandemic”⁵
- ▶ “Low-income or underserved students”⁶

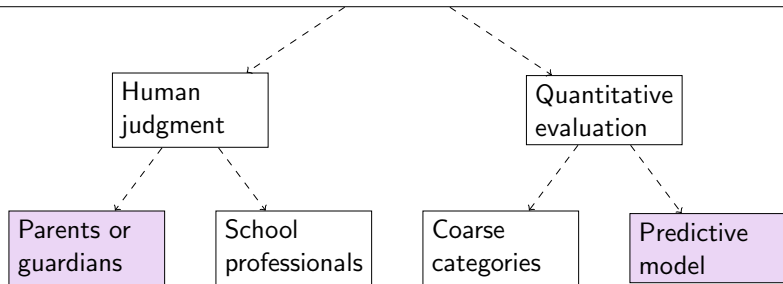
⁴U.S. Department of Education; South Carolina’s ARP Plan

⁵The Hechinger Report; Guilford County, North Carolina

⁶Colorado State House Bill 21-1234

Expertise about need in this case

Targeting statement: “Students who have been most impacted by COVID-19”



Scope of present paper

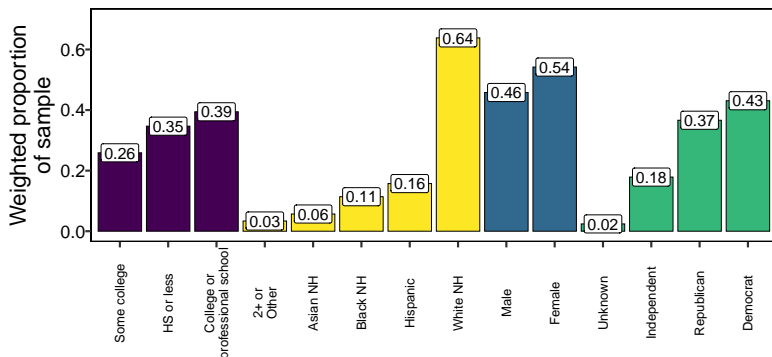
Question	Methodology
What do relevant stakeholders think is the fairest way for judging student need for tutoring?	Vignette-based survey experiment; open-ended responses for underlying reasoning
How are government organizations (SEAs; LEAs) and CBOs judging student need for tutoring? What role are parents playing?	Interviews
What is the causal effect of method T for judging student need for tutoring on which students are offered help?	Simulation; field experiment

Presenting highlighted row; at end, future work aimed at latter two rows

Data

NORC AmeriSpeaks Panel (nationally representative of US adults) with $N = 5,606$ respondents, oversampling current K-12 parents. Fielded in late fall of 2021. Three groups (wording of parent question):

- ▶ **Never parent:** 1,234 ($\sim 22\%$)
- ▶ **Current K-12 parent:** 2,661 ($\sim 48\%$)
- ▶ **Parent but not K-12 (younger or older):** 1,707 ($\sim 30\%$)



Vignette-based survey experiment

A [randomized district racial/ethnic composition] is facing COVID-19 learning losses. While some students are doing fine, others are struggling.

The school district will be giving some struggling 9th-grade students a tutor. The tutor will meet one-on-one with the student during the school day to help the student catch up.

The problem: Unfortunately, one-on-one tutors are expensive. The district only has enough money to give tutors to 15 percent of the many struggling students.

How do schools decide which students get tutoring? Initially, each school's guidance counselor has been: *[randomized to read about one of four counterfactual methods]*

Legal
categories

Parent
requests

Weighted
lottery

Counselor
discretion

Vignette continued...

However, an analytics team within the district has proposed a new method. This method is an algorithm, also known as a predictive model. The predictive model would:

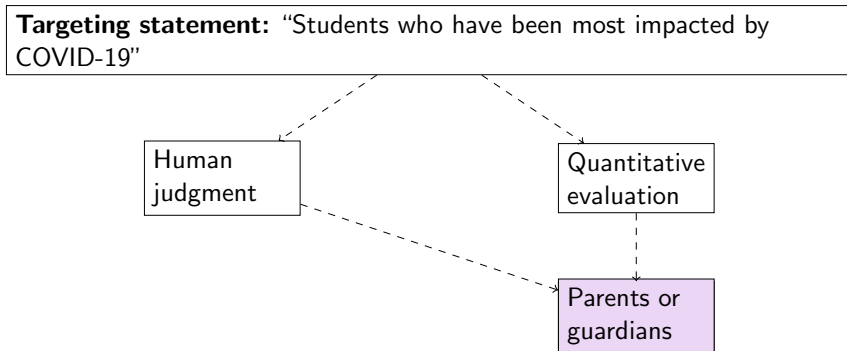
- ▶ Analyze the student records of every student in the district from the past 10 years;
- ▶ Use test scores, grades, attendance, and family financial need to learn what predicts whether a student will fail 9th grade;
- ▶ Then use what it learned to identify current students who are most likely to fail 9th grade.

School counselors would provide tutors to students the model recommends.

Summary: we want your opinion about how the school should give out tutors:

- ▶ **How the school district initially gave tutors:** each school's counselor has been [*randomized other method*]
- ▶ **How the school district could give tutors:** each school's counselor would use an algorithm or predictive model.

Focus here on predictive model versus parents



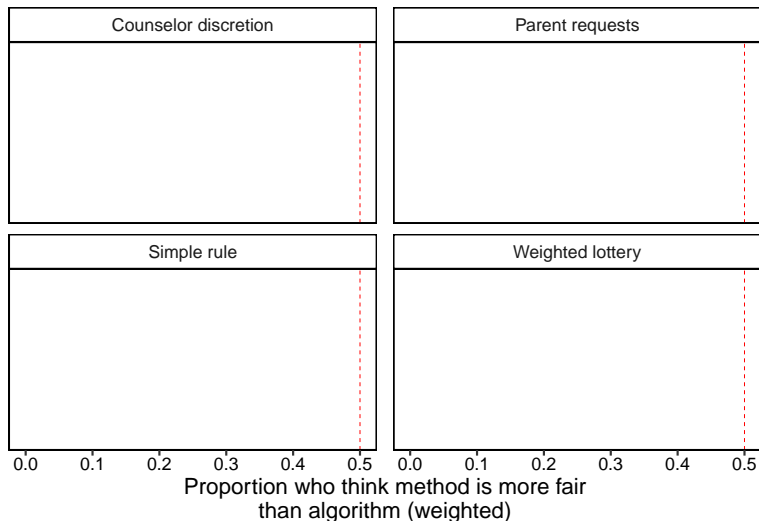
Wording of dependent variables

- ▶ Which method for deciding which students get tutors is fairer?
- ▶ Explain why you think [*method that was chosen*] is fairer than [*method that was not chosen*]
- ▶ When comparing [*counterfactual method*] to the predictive model, how would you rate how certain you are about which is fairer?

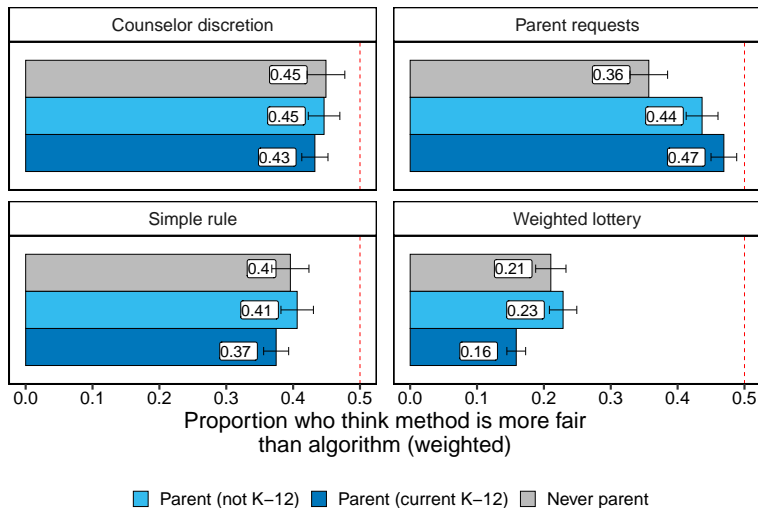
Outline of findings

- ▶ (Unsurprisingly), parents are significantly more likely than non-parents to view parent requests as fairer than an algorithm
- ▶ But overall judgments conceal significant heterogeneity— most notably, political polarization among parents in judgments about *which* is fairer related
- ▶ Open-ended responses reveal different normative views about whether “parent involvement” is a praiseworthy basis for distribution versus a source of inequality

Parents think parent requests are fairer method for ascertaining need



Parents think parent requests are fairer method for ascertaining need than non parents

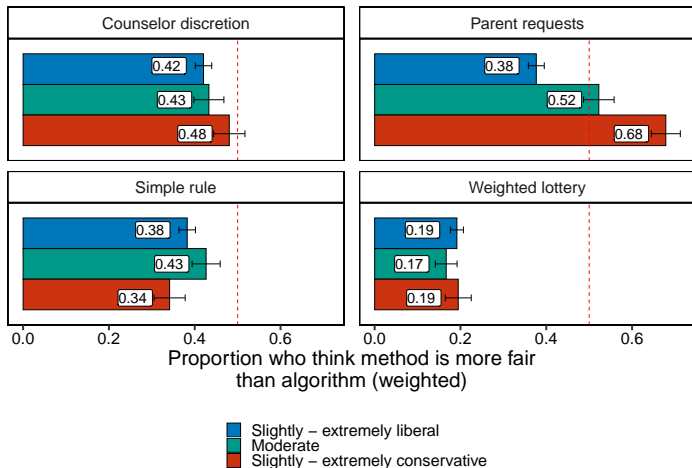


Where we are

- ▶ (Unsurprisingly), parents are significantly more likely than non-parents to view parent requests as fairer than an algorithm
- ▶ **But overall judgments conceal significant heterogeneity — political polarization:**
 - ▶ Quantitative fairness ratings
 - ▶ Qual. responses about why: manual reading
 - ▶ Qual. responses about why: computational text analysis
 - ▶ Interpretation due to correlation with educational attainment?
- ▶ Broader implications and advice on next steps

Conservative parents think parent requests are significantly fairer

Than (1) moderate or liberal parents and (2) the predictive model



What are **conservative parents'** concerns about using **algorithms** to judge need?

Themes:

- ▶ **Parents know their child best**- *similar theme among the liberal parents who think parent requests are fairer*
- ▶ **Tutoring will only work if the receiving student's parents also take responsibility for their education**
- ▶ **Parents' rights as taxpayers**
- ▶ **Distrust of school districts' data**
- ▶ **Suspicion of algorithm as exercise of government power**

White, Female, Slightly conservative, Independent, Some college

Description of algorithm: To judge the current students' needs based on statistics gathered from previous students enrolled.....using financial information which is none of their business and does not necessarily relate to the student's educational progress.

Why parents are fairer: Parents should ALWAYS be involved in any activities relating to their student's education...since it is being provided by their tax dollars

White, Male, Slightly conservative, Republican, College+

Description of algorithm: predictive modeling is a form of socialism and should not be used. Scrap the tutors and hire additional teachers and segregate the weak performers into a separate class where the focus is different.

Why parents are fairer: Parents are the primary persons responsible for education. Predictive models are a step towards socialism where the government decides your fate. That is disastrous in the long run. If parents don't care, the kid has little chance.

Hispanic, Male, Extremely conservative, Independent, Some college

Description of algorithm: Stop using kids as experiments!

Why parents are fairer: Cause they know what the student needs. We know more than teachers and know kids better

White, Female, Slightly conservative, Republican, College+

Description of algorithm: It uses statistics from tracking students in the past years to guess which students would need a tutor.

Why parents are fairer: A predictive model doesn't actually know the current situation of any student. A parents request for a tutor means they are involved and I would assume for outcome of using a tutor is going to have more results than automatically assigning one to certain students.

What are **liberal parents'** concerns about using **parents** to judge need?

Conservative parent theme

Help only works if parents are involved/ take responsibility for their child's education

Example liberal parent concern

"I am a teacher who works in a school where 100% of our children qualify for free or reduced lunches. Our population is 95% students of color. A predictive model is fair because parents who have other concerns as a priority (having enough \$ for basic necessities) rarely are focused on their child's academic needs." (*White, Female, Slightly liberal, Democrat, College+*)

How does the parent's own demographics intersect with concerns about parent requests?

"This model is fair because the students who have means and vocal parents tend to be white (based on experience and research). If the school relies on parent request, more white children will receive support from tutors than black and brown students" (*Black, Female, Liberal, Democrat, College+*)

I believe that a predictive model is fairer than a parent asking for tutor because most parents aren't likely to ask for help due to work or language barrier. (*Hispanic, Female, Liberal, Democrat, High school or less*)

[Algorithms are] a fair unbiased decision making process. Every child will go through the same process the same way (*Asian, Male, Liberal, Democrat, College+*)

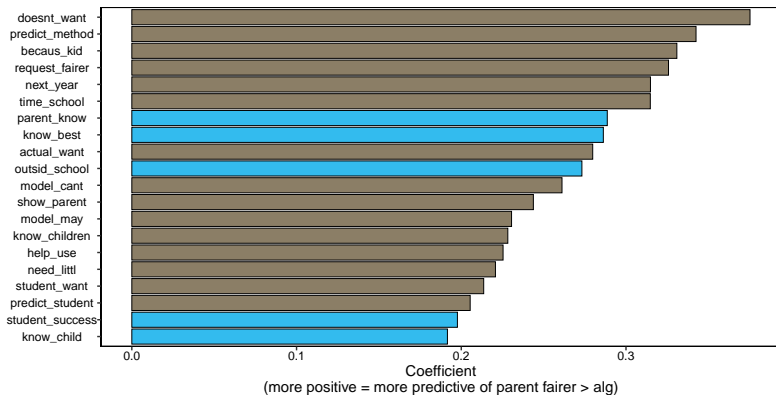
Where we are

- ▶ (Unsurprisingly), parents are significantly more likely than non-parents to view parent requests as fairer than an algorithm
- ▶ **But overall judgments conceal significant heterogeneity — political polarization:**
 - ▶ Quantitative fairness ratings
 - ▶ Qual. responses about why: manual reading
 - ▶ **Qual. responses about why: computational text analysis**
 - ▶ Interpretation due to correlation with educational attainment?
- ▶ Broader implications and advice on next steps

Do these patterns generalize to the corpus?

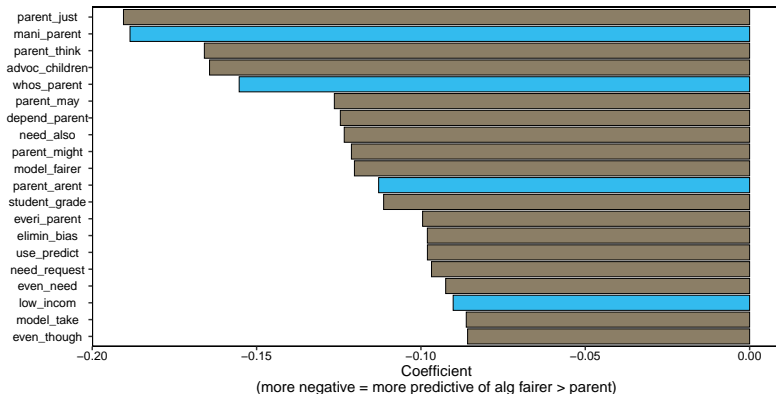
- ▶ **Using penalized logistic regression (LASSO) to select words and phrases predictive of rating parent requests as fairer than algorithm:**
 - ▶ Use lemmatization (paring down words to semantically-shared roots) and stopwords removal
 - ▶ 771 bigrams
 - ▶ [Details on estimation](#)
- ▶ Sentiment scoring using embeddings-derived dictionaries [Results](#)

Bigrams predictive of rating parents as fairer than alg.



More interesting bigrams: go beyond parents knowing child best to utilitarian judgments about student success if selecting on involved parents and outside-of-school factors as reasons to heed parents

Bigrams predictive of rating algorithm as fairer than parents

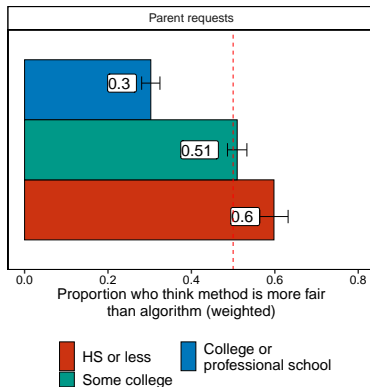
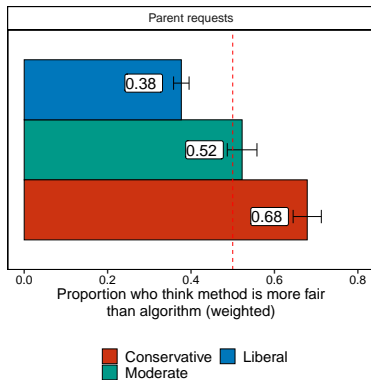


More interesting bigrams: emphasizes inequality between and absent advocacy from parents (mani (“many”); whose parents; low income)

Where we are

- ▶ (Unsurprisingly), parents are significantly more likely than non-parents to view parent requests as fairer than an algorithm
- ▶ **But overall judgments conceal significant heterogeneity — political polarization:**
 - ▶ Quantitative fairness ratings
 - ▶ Qual. responses about why: manual reading
 - ▶ Qual. responses about why: computational text analysis
 - ▶ **Interpretation due to correlation with educational attainment?**
- ▶ Broader implications and advice on next steps

Ideology versus educational attainment

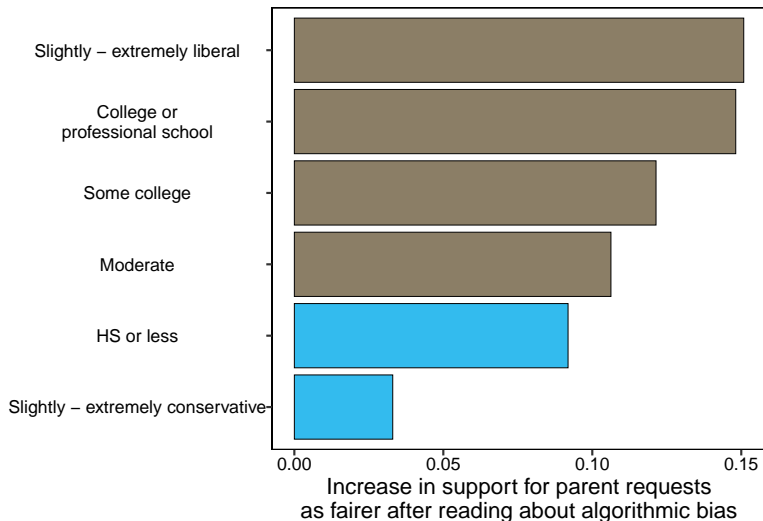


Status update after initial rating

The school district decided to use a predictive model. But a year into using it, the district noticed an alarming pattern. The model worked fine for students who had been in the district since elementary school. But for students whose families moved around a lot, the model incorrectly rated them as low need. That was because the model had no test score data or grades from the students' old districts.

The students who moved around a lot often came from lower-income families. And the model never recommended that these students receive a tutor.

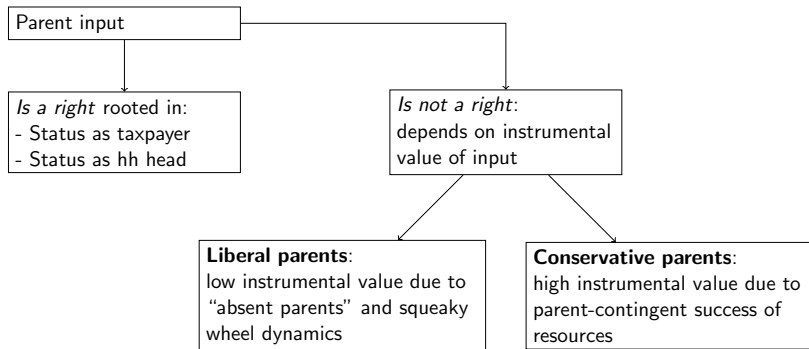
Relative impact of status update



Where we are

- ▶ (Unsurprisingly), parents are significantly more likely than non-parents to view parent requests as fairer than an algorithm
- ▶ But overall judgments conceal significant heterogeneity — political polarization:
 - ▶ Quantitative fairness ratings
 - ▶ Qual. responses about why: manual reading
 - ▶ Qual. responses about why: computational text analysis
 - ▶ Interpretation due to correlation with educational attainment?
- ▶ **Broader implications and advice on next steps**

Normative views on parent input into distribution of goods



Limitations

- ▶ Fielded during a time when parent rights especially salient (09-29-21 VA governor's debate; 11-02-21 election) and insufficient "post election" data to estimate differential trends time trends
- ▶ External validity limitations
 - ▶ **What districts are actually doing:** simple categories (respondents rated these as less fair than algorithm) and little parent input despite federal emphasis on "parent engagement" in ESSER-funded initiatives
 - ▶ **Hybrid methods**
- ▶ Disentangling mechanisms behind support for parent requests
 - ▶ Parents' perception of ideological distance between them and teachers/school officials
 - ▶ Parents' perception of the accuracy of their judgments of need
 - ▶ What right is rooted in (e.g., secular legislation versus natural law⁷)
- ▶ Measure parent's own ideology but not yet partisan context in surrounding school or district

⁷Moschella (2014)

Broader contributions

- ▶ **Parents and K-12 educational stratification:** sociologists tend to highlight negative distributional impacts of parent influence on school decisions;⁸ clash with normative views that emphasize parent involvement as a morally praiseworthy basis for allocation
- ▶ **Organizations and algorithms:** complicate qualitative versus quantitative distinction; multiple sources of qualitative expertise
- ▶ **Polarization and decisions in K-12 districts:** what decisions do parents have rights to influence? Past work on (1) curricular decisions (race; gender ideology) and/or (2) politicized COVID-19 issues (re-opening; masking); spillover onto views about fair distribution of school resources

⁸Cater to opportunity hoarding by higher-SES families (Lareau, 2011; Lewis and Diamond, 2015; Lewis-McCoy 2014; Calarco, 2018, 2020); when parents of color advocate for their child, racialized orgs. mean fewer returns to that advocacy (Lewis-McCoy 2014; Ray, 2019; Manning, 2021; Cartwright, 2022; Penn, 2022)

Advice on next data collection

Book project focus	Additional data collection
Many accounts of ways that algorithms in government exacerbate inequality; in K-12 schools, predictive models have had limited uptake relative to manual points systems/simple categories; understanding why and ethical pros/cons of changes	Interviews with SEA and LEA decision-makers about predictive models versus older methods; ethnography or interviews with edtech vendors; simulation with admin. or other data on “counterfactual fairness”
Polarization around school expertise versus parent expertise; case study of COVID-19 impacts on students	Interviews with current K-12 parents; interviews with SEA and LEA decision-makers; computational analysis of topics of parent rights discussion in news media transcripts versus school board meetings

Appendix slides

Wording of parent question

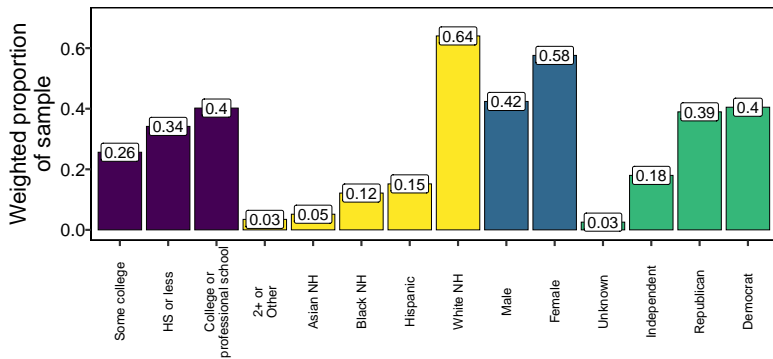
- ▶ I am a parent/guardian and at least one of my children will be in kindergarten through 12th grade during the 2021-2022 school year
- ▶ I am a parent/guardian, but I do not have children in kindergarten through 12th grade during the 2021-2022 school year.
- ▶ I have never been a parent/guardian

[Back to data](#)

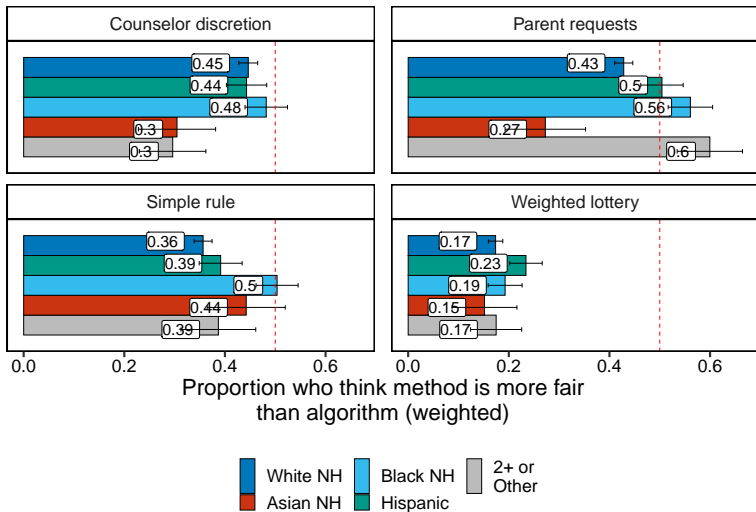
Wording of other conditions

Category Wording Educator discretion	
Lottery	<p>“Initially, each school’s guidance counselor has been using their judgment and personal knowledge of students to decide which students get tutors. The district has encouraged guidance counselors to weigh students’ academic trajectories and family financial need when selecting students.”</p>
Parent requests	<p>“Initially, each school’s guidance counselor has been drawing students’ names randomly to decide which students get tutors, giving some students a better shot at winning the lottery based on their academic trajectories and family financial need.”</p>
Simple rule	<p>“Initially, each school’s guidance counselor has been using parents’ requests as they come in for a tutor for their child to decide which students get tutors. The district has encouraged parents to only request tutoring if they believe that their child’s academic trajectory and their family’s financial need mean the child needs a tutor. ”</p> <p>“Initially, each school’s guidance counselor has been using a test score and family income cutoff set by the school district to decide which students get tutors.”</p>

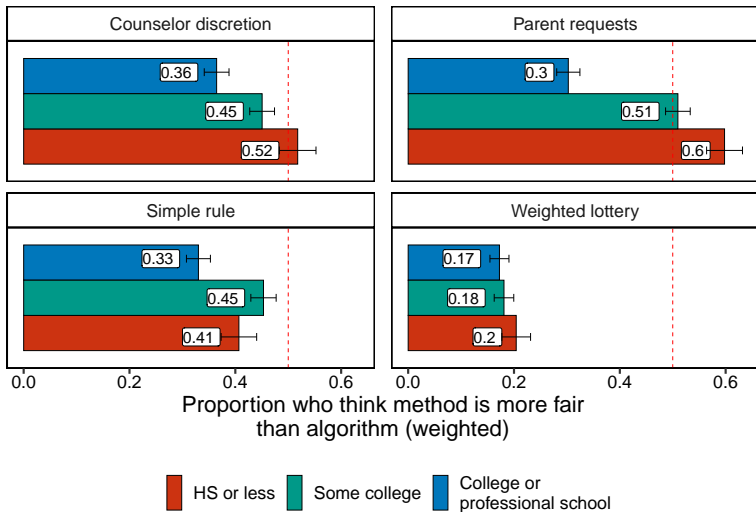
Demographic breakdowns among parents



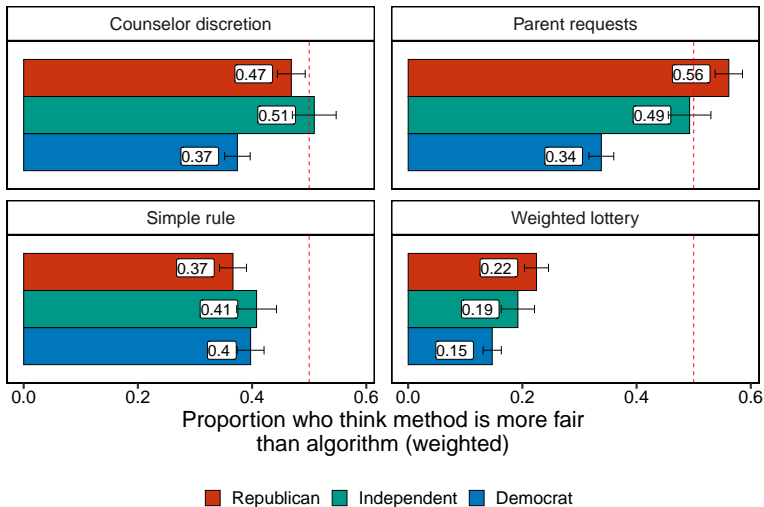
Differences by race/ethnicity among parents



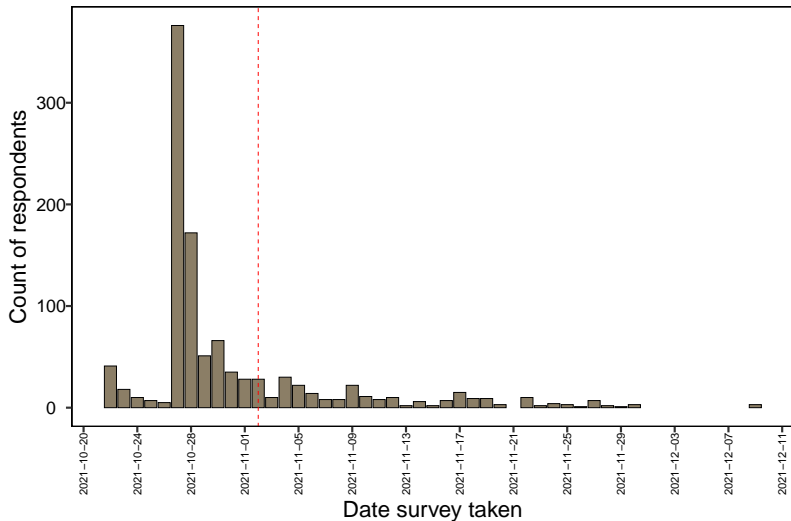
Differences by educational attainment among parents



Differences by political affiliation among parents

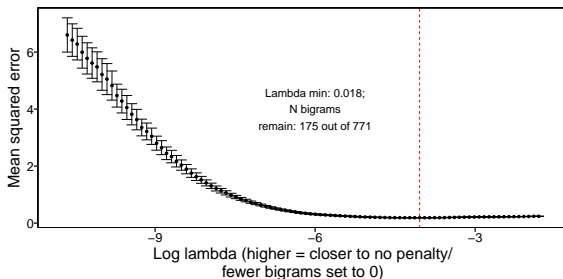


Distribution of responses over time



LASSO predicting fairness rating: details

- ▶ Y : choosing parent requests as fairer than algorithm; X document-term matrix where cols are bigrams
- ▶ `glmnet` in R to estimate
- ▶ λ chosen via cross-validation
- ▶ MSE-minimizing λ : 0.018—relatively weak penalization
- ▶ **Next steps:** add demographic features and compare bigram-only prediction versus bigram + dem prediction



Embeddings methodology

- ▶ Used pre-trained embeddings derived using the Global Vectors for Word Representation (GloVe) method (Pennington, Socher, and Manning, 2014) on the corpus of 2 billion tweets ([source](#)); used tweets rather than Common Crawl (general web scraping) or Wikipedia due to similarity between informal language in free responses and tweets
- ▶ Seed words reflecting low instrumental value of parent input: advantage; bias; afford; poor; time
- ▶ Procedure⁹:
 1. Look at overlap between seed words and words in responses (all overlap)
 2. Estimate the similarity between the vector of seed words and words in the embeddings corpus (cosine similarity with $L2$ norm); use top-k most similar words ($k = 200$ for present results) to build a larger dictionary corresponding to that theme/sentiment
 3. Calculate a given response's weighted overlap with the dictionary (e.g., if word A with similarity weight 0.8 appears once, contributes $1 \times (1 + 0.8) = 1.8$ to score; if word B with similarity weight 0.4 appears once, contributes $1 \times (1 + 0.4) = 1.4$ to score; numerator for sentiment score)
 4. Standardize by count of total terms (most don't overlap with dictionary) as denominator

⁹Adapted from Choi, Harris, and Shen-Bayh, 2022

Top responses: mixed coherence of theme

Response
because they can see how the students function not just what parents think
i really dont know what that means. will like to know to give better answer
parents know their children and know when they are struggling. think both ways work they just need to be modified
i think its going to offend some students and parents to be categorized
because thats how they will tell who really needs the support.
they know whats going on?the parents just think they know tutoring cant work without cooperative parents and students.
every student will be assessed that fits the requirements.
i feel that more students need tutoring than the parents think/feel their children do.
parents are going to know if there child is in need. models are great but parent knows for sure if there child needs help. spend the money on getting more people vs money spent on the predictive model.