



OZAN JAQUETTE <ozanj@g.ucla.edu>

Fw: [EXT]EEPA - Decision on Manuscript ID EEPA-22-SIP-4256.R1

Salazar, Karina G - (ksalazar) <ksalazar@arizona.edu>
To: OZAN JAQUETTE <ozanj@ucla.edu>

Thu, Feb 16, 2023 at 5:12 PM

From: EEPA <onbehalf@manuscriptcentral.com>
Sent: Thursday, February 16, 2023 5:52 PM
To: EEPAeditors@aera.net <EEPAeditors@aera.net>
Subject: [EXT]EEPA - Decision on Manuscript ID EEPA-22-SIP-4256.R1

External Email

16-Feb-2023

Dear Author,

Your manuscript # EEPA-22-SIP-4256.R1 entitled "A sociological analysis of structural racism in student list products" which you submitted to EEPA for the special issue on Critical Approaches to Educational Policy Research has now been reviewed. Based on the reviewers' recommendations, we cannot accept your paper in its present form. However, we are interested in receiving a revised version of the paper that addresses the reviewers' concerns.

The reviewers agree that the piece has the potential to make a significant contribution to the policy sphere, particularly the focus on an understudied part of the college admissions process. This piece has the potential to expand our understanding of the (racialized) tools used by colleges to create their incoming classes in ways that help illuminate equity concerns and lay out guidelines for future scholarship.

Despite the promise of this work, there are several significant concerns raised by reviewers that need to be addressed for this manuscript to be considered in EEPA's special issue. We highlight several key areas to focus on below, but we encourage you to respond to and address each of the reviewers' comments. It will take a significant lift for revisions of this paper to demonstrate a compelling contribution to the field. We believe that this is possible, but want to make clear that this would be a substantial revision. We have organized the suggestions into major and minor categories.

Major

First, Reviewer 1 and 2 note that several of the arguments made in the paper are not necessarily about college lists, but instead about the admissions priorities of each institution. We believe it would be useful to tease this out more in the paper. We also believe it would be helpful to provide some (masked) information about the 14 institutions. This could be rounded enrollment numbers (in bands) or Carnegie Classification or something else. But it feels like as the argument is strengthened on why this might be about the admissions priorities of institutions, the reader would benefit from better understanding the contexts of the 14 institutions.

Second, Reviewer 1 notes that the current analysis does not provide a substantial contribution to our understanding of admissions recruitment. The reviewer provides suggestions of ways the authors could expand the analysis to provide substantial insight into recruitment. We are not being prescriptive to say that the authors must follow Reviewer 1's guidance. However, there needs to be more done. The current findings generally appear to align with what we already know about who takes college entrance exams or has higher academic GPAs. The current tests of proportions are not that revelatory given that the appendix tables show that almost every estimate is statistically significant (likely partially due to Type I error but also because these are large samples that will find practically insignificant differences statistically significant). More interesting are the findings that combining these elements creates a substantially more exclusive prospect pool (more could be done to explicitly tease this out).

This also applies more specifically to the discussion section. Arguments are put forward that "several academic filters and geographic filters are structurally racist inputs" yet there is no clear explanation of when an academic filter would be a structurally racist input and when one would not be. For example, if an institution actually did use GPA and zipcode to expand the diversity of the prospect pool, is that a structurally racist input? If so, can structurally racist inputs be used for good? These are the more interesting, and challenging questions, that the paper could engage in. This would then also allow the authors to provide clearer policy recommendations.

Third, Reviewer 2 and 3 both seek more of an overview in the paper on college admission and recruitment broadly. We understand that this is a sociological analysis, but it does feel like somewhere in the introduction or literature review, it would be helpful to include some space devoted to an overall overview that cites interdisciplinary research on college recruitment. Given that EEPA covers all levels of education and all disciplines, it would be useful to ensure that readers who are not experts in this individual area can have strong grounding in how college recruitment works.

Fourth, we agree with Reviewer 1 that it would be helpful to include citations or peer-reviewed references to justify assertions throughout the paper. We can certainly see why some statements are likely to be true, but the paper needs to provide clear textual citations that provide the evidence behind the statements. As an example, across pages 5 and 6, there's a statement about the test-optional movement being an existential crisis for the College Board. Yet, the College Board makes the majority of its funding from administering the AP program (and has expanded its revenue streams by contracting with states to provide the SAT as a high school exit exam). If the authors want to make that argument, there needs to be a clear explanation of the evidence that drives the argument. Another example is that the authors repeatedly argue that micro-targeting is used for harm but, as Reviewer 1 notes, there needs to be justification for this argument included in the paper (this ties into the next point on the conceptual framework). These types of statements can be found throughout the paper. Therefore, a thorough revision of the entire paper attending to ensuring there's justification for statements of fact would be beneficial.

Fifth, while well written, the conceptual framework seeks to cover several different areas of scholarship. In line with trimming some of the front matter of the paper to fit page limits and to provide a more coherent narrative, it would be useful if the authors selected one to two of the theories currently described. For example, the authors might focus on critical data studies to examine how algorithms are constructed/trained and the ways institutions might wish to use them for good but could ultimately create more harm. This shift would allow more space to provide the additional justification for the authors' arguments.

Sixth, based on our read of the article, the ties to education policy need to be made clearer. We appreciate the discussion of implications on federal regulations. To create a stronger connection with policy, we recommend the authors look at the feedback from Reviewer 1. We also noted that the discussion of policy relevance solely focuses on the federal government. Yet, these are all public institution student lists. It seems that more attention needs to be given to the role that state policy actors have in ensuring institutions are using these lists in an accountable way. This would likely necessitate a bit more attention to the fact that these institutions are public which likely means they also have higher acceptance rates than their private peers and different ways of marketing & evaluating applicants. For example, the last sentence on page 6 notes the private school counselors can do a lot to get students into "top colleges" by communicating that they will definitely enroll at that institution. Yet, this type of demonstrated interest is frequently not used at public institutions. It would also be helpful to say more about whether states or the federal government have considered any policies or new regulations about these lists.

Seventh, we have concerns about the geographic analysis. HSLS is a dataset that is representative at the national level and at the regional level. It is not representative at the city or zipcode level. That means that, even with weighting, all the geographic analysis is on a unique sample (that does aggregate up to being regionally and nationally representative but is not reflective of these smaller geographic zones). While we're not saying these analyses must be removed, the authors need to seriously grapple with what the geographic analyses can provide evidence for. If the authors decide to leave this analysis in, there must be significant justification for its inclusion and discussion of just what these random data points can tell the reader about their underlying geographic areas.

Minor

Eighth, given that the SAT and PSAT variables are actually measures of SAT/ACT (in SAT units) and PSAT/preACT (in PSAT units), it might be helpful to name those something like "college entrance exam" or "pre college exam" or something when talking about those thresholds. As currently written, the paper can be confusing to the reader about what those variables actually mean.

Ninth, Table 1 would benefit from a note that details the data (e.g., the fact that it covers multiple years, what a geomarket and segment is).

Finally, we recognize we have requested several spaces where the paper needs to engage more on topics. Given the page length concerns, it feels necessary for the literature review and discussion sections to be streamlined a bit more (the reviewers provide some suggestions for reorganization and trimming) and it also feels that Figure 1 (which is a reproduction of other scholars' work), Figure 9, and Figure 10 could be cut or moved to the appendix. We agree with the reviewers that the authors may not have the legal expertise to provide some of the suggestions they do so it would likely make sense to remove the top paragraph on page 32. The first full paragraph on page 33 can likely be removed as well since it gets very detailed on certain parts of platform that are more ancillary to the larger point of the paper (it also includes a sentence that appears to be the same as a different sentence in the introduction). Several paragraphs on page 34 can likely be trimmed to a single sentence. These are just some suggestions for spaces but the entire manuscript should be evaluated to see where the narrative can be made tighter.

In recognition of the aim of this special issue to further conversation, understanding about, and attention to critical approaches to policy analysis in education, we are also requesting that all authors attend to a few additional points. Terminology varies across fields and terms common in one area may seem like "jargon" to others. Please take care to define key terms and unpack theories for this generalist audience. Provide detail on your data and methods, including explanation and justification of your analytic choices and processes (e.g. provide examples of codes), especially for qualitative and discourse methods that are less common in this journal. Address the implications of this work for policy and/or research. We hope that readers will use this special issue to reflect back on their own work, even if it is not in the same substantive domain of policy research. We believe that these measures will help strengthen the special issue and the role it can play in the broader field of education policy research; we appreciate your help in accomplishing these aims.

Below please find the reviewer comments regarding your paper. If you disagree with any of the reviewers' comments or feel that any of their suggestions are unreasonable, please submit a written justification for your position along with the revised version of the paper. A point-by-point reply to each of the reviewers would be very helpful.

We would like to receive your revision by May 1st and we will send the paper out for external review at that time.

If extenuating circumstances do not allow you to submit your revision by the due date, please let us know immediately. If we do not hear from you and the deadline passes before the manuscript is submitted, the paper may not be included with this special issue and may be considered as a new submission.

Please understand that EEPA has space to publish just a fraction of the manuscripts submitted. The current acceptance rate is less than 10 percent.

We truly appreciate the opportunity to review your manuscript, and hope that you will continue to think of EEPA for your future research. We wish you the best in your continuing research.

Sincerely,

Dominique Baker
Guest editor, EEPA Special Issue on Critical Approaches to Education Policy Research
djbaker@smu.edu

Geoffrey Borman
Editor, Educational Evaluation and Policy Analysis
EEPAeditors@aera.net

Reviewers' Comments to Author:

Reviewer: 1

Comments to the Author

In this manuscript the authors work to demonstrate that the structural racism underlying the search filters used to generate student lists for universities to use in their recruitment efforts causes the lists they produce to systematically exclude groups of students who are traditionally underrepresented in higher education in the U.S. They do this by simulating the results of various search criteria using the nationally representative data in HSLs. In each simulation, the student lists generated by the search criteria are less diverse than the population of students overall. The authors suggest in the end that the use of student list services may warrant more legal scrutiny.

Overall, I applaud the work the authors have done to obtain information about the student list requests submitted by public colleges. I agree with them that this part of the college recruitment process is under-studied and under-scrutinized—there are almost certainly problematic aspects of the construction and use of student lists. The argument overall was thought provoking—reviewing this manuscript made me think deeply about my own biases related to parts of the admissions process I largely take as given.

I am not fully convinced by the argument as it stands, however. Some of this may be an issue of my misunderstanding, in which case I hope the authors use my blundering as a signal for whether they could work to clarify their argument. I also want to push the authors to find ways to make the analysis feel more substantial. As I outline below, the bold setup for the paper left me underwhelmed by some of the findings, which felt at times already well known or poorly motivated.

Finally, I'd strongly caution against the legal speculation here—it rests on quite a set of logical chains that I don't think have much legal basis. I don't think it is necessary for the manuscript to go in this direction.

I start with some of the big questions I have, which suggest broad areas where the authors may want to push their (or, clarify my) thinking, and then move on to more specific thoughts and opportunities to improve.

Where does the problem lie?

The manuscript, in large part, seems inspired by the earlier FOIA request of student list purchases and the filters that determined the content of those requests. This is a cool FOIA request to think to make, and obviously a lot of work. The nature of the request made me think we're dealing

with how universities use student list requests either implicitly or explicitly to privilege certain groups in the search and recruitment process.

This understanding of the purpose of the manuscript is dismissed a few times, however, in part by the assertion that universities, despite prior assumptions, do not do recruiting “individually,” but rather alongside third party enrollment management consultants. Here I wondered whether these third-party businesses would be the focus of the work, but these organizations largely stay on the periphery of the argument. In the end, I’m confused how much independence these businesses have from universities—don’t they act on behalf of the university’s stated enrollment preferences (which would still make me think that universities are independent when determining recruitment priorities)? Or is the idea that they engage in student list filtering in a way that is more nuanced than what universities would do on their own and that universities may not understand the students who are excluded by these filters? Either way, we do not observe any actions by third-parties in this research (which makes me question whether they need as many words devoted to them as they do). The manuscript returns to third parties in the conclusion, specifically by citing that the US Department of Education thinks that the third party enrollment goals are determined by institutions—the implication is that this may not be the case, but no evidence is presented to support this.

Instead, the manuscript turns to algorithmic customer selection tools. A lot of baggage is placed on the word “algorithm,” but I worry about this because its definition and application to the student list purchasing process is a bit of a moving target. Initially, algorithms are simply any computer executed code. This definition doesn’t seem all that analytically interesting, but then the manuscript layers on actuarial practices where we try to use past behavior to predict future behavior. This does seem problematic in the case of college admissions because if we assume students always behave as they have in the past, then clearly we can’t use these algorithms to include previously excluded groups. Where I get hung up is that I don’t see this level of prediction in the student list filters, but maybe it just needs to be made more explicit for me. I certainly see it, for example, in the College Board’s geodemographic segments that can filter on past college-going behaviors, but again we don’t observe either that prediction or even those filters in this research.

The ultimate question appears to focus just on whether the filters themselves are structurally problematic. I think this may be a reasonable question, except that we don’t need the research here to know the answer—even before the analysis, the manuscript cites literature that shows that gaps in test taking and scores by student background characteristics which necessarily suggests that filters using those criteria would exclude certain groups. Is there doubt that these patterns have persisted since that research was produced?

How are geographic filters used?

My impression here has always been that a university could say, “I want students from these states/counties/zip codes...[and then list them].” I imagine this option is available for public universities to target students in their home (or neighboring) state or within certain catchment areas (like the Cal State campuses have), or for private universities to recruit knowing that students tend to like attending college close to home. It would be helpful to confirm (and maybe you have) that selecting geographic areas by income characteristics is even an option (though I suppose it’s possible a university could find those zip codes on their own and then ask for them specifically in their list).

Even if it is a common way of conducting geographic searches, I’m a little incredulous that a university would specifically search on geographic area income levels. I know that Author XXXX(d, e and f?) found that off-campus recruiting tended to focus on wealthier schools in larger metropolitan and suburban neighborhoods, but I don’t think it’s safe to assume that schools would pursue the same recruitment strategy in both the student list development and their off-campus recruiting, or that this focus is the only recruitment priority of universities in their recruitment. As the manuscript notes, off-campus recruiting is often used to maintain relationships with certain high schools—something that being in person facilitates. Given the well-worn path to reaching students in these schools, it doesn’t make sense for universities to duplicate that effort when buying student lists. Maybe they do it in some cases, but it seems like a lot to base a whole section of analysis on speculation that colleges “may” do something.

I reached out to my own admissions office and they use their student list purchases for two purposes. First, they use it to find first-generation, rural, and racially minoritized students they don’t find through other modes of outreach. Second, they use it to complete details about students (e.g., their high school) who they may have met at a college fair for whom they might only have a name and address. I know a former institution

of mine similarly uses search to find students of color, and first-gen students. I'm not sure either of these admission offices would recognize the sort of searches you suggest with respect to looking for high income zip codes, though this is obviously based on anecdotal data with an n of 2. Either way, this part of the analysis feels unsubstantiated.

When does the student list fail to create access?

I certainly agree that structurally racist systems should not deny any student the opportunity to pursue a postsecondary degree. Thus, the fact that the filters that generate lists used in college recruitment include fundamental racial bias is a problem. However, I think it's fair to say that the inclusion of, for example, academic filters in search criteria exist because of the centrality these academic factors have in admissions decisions themselves. Neither students nor universities benefit if a university put resources into recruiting students who are unlikely to be admitted to that university. Thus, it may be worth exploring whether student list generation is the problem, or the priorities of the admissions system in general. It is also possible that universities don't cast as wide a net as they should—if you know the admissions statistics of the universities that you FOIA'd, as well as the thresholds those universities use in their search criteria, you could figure out whether universities were genuinely trying to capture all reasonable applicants (I'd hope for searching for students at their 25th SAT percentile or lower), or if they set their thresholds too high to be maximally inclusive.

This also makes me wonder whether “all other students” is the right comparison for included students. Regardless of their qualifications for selective admissions, not all students have an interest or desire in attending college. We obviously want to maximize the group who does what to go to college to attend, but I worry about whether you stack the deck too much in favor of your argument by including everyone in the excluded group descriptions. At the very least, I'm curious the extent to which students eligible to show up in a College Board search—they have a PSAT, SAT, or AP score—look like the population of students who enroll in college at all (I think it might be interesting here to think about both four-year college enrollment, and any postsecondary enrollment). If the college population is more diverse, it tells us that colleges are able to find more diverse students than they would if they relied on lists alone and that lists would fall short in helping universities identify appropriate applicants. If the list-eligible students are more diverse, it emphasizes the problematic elements of college recruiting in admissions processes.

Perhaps adding columns to the existing results that describes the composition of excluded students who ultimately enroll in a four-year college or any college at all would add strength to the analysis. I recognize the problem of assuming the demographics of the contemporary college-enrolled population as the “right” composition to strive for, but I suspect the list criteria would still have a hard time meeting this benchmark as a *minimum*.

More concrete thoughts...

Do we have any sense of what percentage of postsecondary institutions buy student lists? Is it prevalent everywhere? Just four-year colleges? Just selective colleges?

p. 1 How much weight should we put in the study about students opting in/out of the College Board's search tool? Aren't students taking pretty clear stances on their orientation to college enrollment here? From the description in the manuscript, it sounds like this is a descriptive study with some pretty big omitted variable bias...

I might work to refine the research question on p. 2., which asks “what is the relationship between student list search filters and the racial composition of students who are included versus excluded in student lists purchased from College Board?” I don't think the data can answer this question (particularly since you don't appear to have data on specifically how the filters are used in the requests from the College Board).

p. 4 I have a hard time parsing the final sentence before section 2.2. breaking it up in two sentences may be helpful. Also, if footnote 2 provides results for private institutions, what universities are included in the main text?

p. 5 I'm not sure I fully followed, or found much value in the timeline of search firms—a lot of the names, for example, won't mean much to the average reader. I also had trouble following where the competition and concentration was in that paragraph. Revisit this section to see how much

is necessary for the overall argument.

p. 6. In the sentence about off campus visits and enrollment priorities, be clear about whose enrollment priorities. Also, is there another citation to use here? It's a big claim and it would be good if someone other than you sees this link.

p.7 Is there a citation for the claim that universities are uninformed about which students they target? (last paragraph of the page)

p. 11 In the first full paragraph, the manuscripts indicates that it is important to understand why "birds of a feather flock together" as the College Board states. It would be good to finish this thought and say why. If colleges are looking for diversity (which, maybe I naively think some of them are—at least a few of them seem to want to continue to use race conscious affirmative action, for example), and demographic sorting allows them to do it, explain why the cause of the sorting matters from the perspective of a diversity-seeking college. I'm not sure all readers will get there themselves. The manuscript returns to this idea on p. 15 and I'm not sure a reader would follow the logic any easier there.

p. 12 Be careful with your argument about the critique that "classifications systems are developed to optimize profit." That may be true in the business world, but I think many readers would argue that colleges don't set out to optimize other things than (at least immediate) profit. Does that change the conclusions we draw from this claim?

p. 17 Can you describe the 830 College Board orders in more depth? I'm still not sure what data you actually have here—is it what filters were used (broad categories) or what levels of those filters (e.g., SAT scores above 1200...). How do these data enter the analysis?

p. 18 Explain whether X3TGPAACAD is weighted or not and how this aligns with what students might report to the College Board.

p.23 Would the College Board have GPAs for anyone who didn't take the PSAT, SAT, or an AP exam? Is that exercise in the analysis just a question of "who took any of those exams?"

p.23 Don't forget your CITE in the footnote.

Throughout the results: I don't know what $p < 0.000$ is as a significance level. Is this just a rounding/significant digit thing based on your statistical software? If so, report the lowest reasonable p-value that this rounding infers (I'm guessing $p < 0.001$)

p. 30 (and other places) I think there's a substantive difference between filtering by zip code (e.g., "we want students in these zip codes in the catchment area for CSU Long Beach") and filtering on zip code affluence, but often filtering on affluence is referred to a simply filtering by zip code, which I find confusing at best. Keep the distinction clear.

Other questions/things to think about:

How should we think about inclusion and exclusion based on SAT scores in a world where more states and localities are now requiring high schools to take the SAT/ACT?

How should a reader think about the timing of student list purchases? For example, purchases based on PSAT score allows universities to reach out to students much earlier in their college search process, potentially allowing them to shape not just the students' preferences for college, but how they prepare. Purchases based on AP scores or SATs likely can only sway marginal preferences for where students apply.

Products like the College Board's search tool are blunt and imperfect and, as you demonstrate, likely problematic and exclusionary. But there aren't many alternatives. What would an ideal tool look like that allows colleges to find students (affluent and otherwise) that have a reasonable likelihood of being admitted to the institution?

Reviewer: 2

Comments to the Author

The paper is an important and critical look at an understudied area of college access and admissions literature—the recruiting process. The authors draw on nationally representative data paired with actual criteria used by colleges to generate and purchase student lists to demonstrate the racialized process through which colleges recruit prospective students. It's a fascinating inside view of how common admissions practices that draw on racialized inputs result in inequities in college recruiting (and subsequently, college access). I appreciate the care the authors took in collecting the data (3 years!) and creativity in applying the filters commonly used by colleges to nationally representative data. The results are sadly not surprising and offer clear evidence in support of the authors' propositions.

I offer a few suggestions that I think would help clarify the importance of this study:

The first is to step back a bit and provide a description of college recruitment more broadly: what colleges participate in recruiting? What is the motivation for colleges to recruit students and how does this differ depending on the type of college (e.g., public versus private, highly rejective versus open access)?

Related, I think it would help to specify just how important student lists are to colleges. Perhaps this could be accomplished by clarifying what colleges use student lists (is this mostly highly rejectives or are regional comprehensives, community colleges, and other less rejective institutions using student lists to recruit?) – this helps identify the particular colleges that are more or less likely to engage in this process and the extent to which student lists are used in recruitment. Additionally, perhaps some of the authors' prior work highlight just how important these lists are to colleges? For example, are they the primary way colleges identify a pool of students to recruit (versus looking at feeder high schools and prior year stats?)? I think it would help to discuss just how important these lists are or are not and to what institutions.

The conceptual framework is well-written and does a good job connecting search filters used by colleges to structurally racist inputs. One thing that I think would strengthen this section is to offer a more detailed discussion of how these inputs come to be racialized. For example, the authors note that zip codes are racialized but do not offer citations or a deeper discussion of this. Offering examples of policies and practices that have led zip codes to become a racist input would help strengthen the paper's contribution. Similarly, the authors note factors that contribute to inequities in test scores by race, but this discussion, I think, is fairly surface level and doesn't fully engage with the extent to which resources have been extended to some, primarily White schools, and denied to other schools in ways that create different educational opportunities. The authors start to do this (e.g., on page #), but I think a more detailed discussion would be helpful in establishing just the extent to which these inputs are likely to be racialized. I could see this being a great piece to cite in the broader admissions and financial aid literature (regarding racialized evaluation metrics), so more developed connections in the conceptual framework would be useful.

One point that seems to be missing in the discussion is that the reason colleges select on filters in the first place is that they are often selecting on the same set of filters that they use to evaluate admissions applications. As a result, we could (and should, as the authors clearly show) change the way colleges recruit students, but absent additional overhauls to admissions criteria (and financial aid, college culture, etc.), the same outcomes are likely. Maybe some way to build this into the discussion is to highlight whether implications from this study for recruiting could similarly be used to rethink admissions criteria. I defer to the authors for how they want to address this, but it seems important to note that the entire system (K12 education, college admissions and aid) is designed to produce these inequitable outcomes, so changing recruiting practices is one lever, but other levers similarly need to be pulled to produce different outcomes.

Minor notes:

The authors mention "platform studies in higher education" a few times in the manuscript – it'd be helpful to define this at first mention and discuss what some of these platforms are and the role they play for universities

The second paragraph notes that student lists are important for degree completion – it wasn't immediately clear to me what the connection here

is, so making it explicit could help.

The first full paragraph on page 2 offers a lot of key definitions for the paper one after another – it's easy to get a little lost in this paragraph, so maybe it would help to offer more transitions or descriptions of each key term and clarify how it connects to others

The end of the introduction methods unfair practices criteria under the FTC Act, but this is the first time it's mentioned, so it would help to clarify what the authors mean by this and how it relates to the student lists (the authors are clear about it in the discussion, but the first mention doesn't have much context)

Is there a citation for Figure 1? Or does it draw on previous literature to identify the funnel?

The authors mention a period of vertical acquisitions in the enrollment management industry, but I'm not sure I saw a discussion of these vertical acquisitions in the text that follows

Page 5 – authors mention ACT and College Board offer clients information about prospects that are not included in student lists – what type of information is this?

In section 3 that starts on page 6, the authors mention prior scholarship that focuses on the later stages of the enrollment funnel – perhaps cite some of this literature?

“Blind spot” on page 6 – consider changing the language to remove the word blind

On page 8, the authors mention critical geography, but I think this is one of the first times it is mentioned in the manuscript – maybe mention this in the introduction and offer a brief description of how it informs the study

Page 9 – what is the connection between actuarial methods and selection devices? I think making this explicit would be helpful for the reader

In section 4.2.1, the authors posit two broad sources of structural inequality – might be worth noting that these two sources are deeply connected – i.e., the input itself determines who is included/excluded in the underlying data while the search filters further exacerbate these differences in who is included/excluded

Round sample sizes to nearest 10 for HLS09 data reporting throughout

In discussing the 2 data sources on pages 16-17, can you clarify which data are used to investigate each proposition?

Are both Table 1 and Figure 3 necessary? It might be possible to just have one of these given the number of tables and figures

In 6.3 starting on page 27, it might be useful to reference back to table 1 or figure 3 to note how these combined filters were selected (i.e., based on common search filters)

The authors note how much colleges pay per name, but it might help to provide an overall amount some colleges spend on student lists, just to give the reader a sense of the importance these lists play in recruitment

Reviewer: 3

Comments to the Author

In this paper, the authors introduce readers to the current practice of student list-buying and argue the structure of these lists systematically exclude students of color in two ways: (1) the lists historically exclude non-test takers and (2) research filters used to control which prospect files selected upon structurally racist inputs. The authors use HLS09 to recreate the College Board Student Search, testing a number of propositions about the relationship between search filters and racial exclusion of students from historically underrepresented communities of color. Their descriptive analysis is grounded in a strong and clear review of research and scholarship from sociology, actuarialism, and critical data studies. I very much enjoyed reading this paper and hope that the suggestions and considerations below will strengthen it even more.

Suggestions and considerations:

I would suggest the authors consider situating this study in a broader literature around college access. There is one paragraph in the beginning that discusses the probability of enrollment for students opting into the College Board Search Service, but I think there is more that could and should be said about existing racial inequities in college enrollment. What are current figures on the representation of students in public universities, for example? Have these numbers changed in recent years? Are they likely to worsen given affirmative action legislation or improve given the movement toward test-optional policies? In short, a brief discussion on the current representation of underrepresented groups in higher education helps to contextualize the broader impacts and implications of student lists and situate the magnitude of the problem that the study is

illuminating. In essence, the so what? Likewise, the paper should end by coming back to this point about equitable access and opportunity to college for students from minoritized backgrounds.

The systematic exclusion of applicants of color is a key argument in the paper. As such, the authors might consider discussing the Hirschman and Bosk piece in more detail. It is cited throughout and seems to provide critical support for the authors' claims. Is the current conceptual framework an extension of or a response to Hirschman and Bosk's work?

The mention of micro-targeting and market segmentation made me think about critical quant studies in education policy research and the work of scholars (e.g., Samantha Viano and Dominique Baker) raising questions about the construction of large-scale datasets such as what are the types of variables that are being included. How are race and class conceptualized? How are other critical variables operationalized? It seems this literature is quite relevant to the paper and I would encourage the authors to consider featuring it more prominently or at least citing it in the discussion of critical data studies.

I was unclear/a bit confused by section 4.2 of the manuscript. It seems to repeat the preceding discussion of algorithmic selection devices. Is this the start of the conceptual framework? Should these mechanisms be collapsed in or incorporated into section 4.1 which discusses them as well (although using examples from outside of education)?

In section 4.2.1., there is a sentence, "Drawing from theory..." I would just add in theory for a better description. For example, sociological theory.

I would encourage more organization of section 7, the discussion of the paper. At times, this section is a bit wandering and repetitive. Consider adding subsections for "Implications for Research" as well as subheadings to organize the various implications/future directions for research. And as stated above, I would circle back to the broader implications of this work for college access.

Instructions for Resubmission:

To revise your manuscript, log into <https://mc.manuscriptcentral.com/eepea> and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. It can be helpful for reviewers if you use colored text or the yellow highlighter to show where significant amounts of text are new or changed.

Once you have revised the manuscript, you can upload it and submit it through your Author Center.

You also now should respond to the reviewers' comments in the space provided. Use this space to document your responses point-by-point. Please be specific in your response to the reviewers.

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

So that we can continue to publish manuscripts as expeditiously as possible, please submit your revision as soon as it is complete. If you are not able to meet the deadline for the revision, please let us know as soon as possible.