

# BROOKINGS

Report

## **Enrollment algorithms are contributing to the crises of higher education**

Alex Engler Tuesday, September 14, 2021

### **Editor's Note:**

*This report from The Brookings Institution's Artificial Intelligence and Emerging Technology (AIET) Initiative is part of "[AI Governance](#)," a series that identifies key governance and norm issues related to AI and proposes policy remedies to address the complex challenges associated with emerging technologies.*

## **Executive Summary**

**H**undreds of higher education institutions are procuring algorithms that strategically allocate scholarships to convince more students to enroll. In doing so, these enrollment management algorithms help colleges vary the cost of attendance to students' willingness to pay, a crucial aspect of competition in the higher education market. This paper elaborates on the specific two-stage process by which these algorithms first predict how likely prospective students are to enroll, and second help decide how to disburse scholarships to convince more of those prospective students to attend the college. These algorithms are valuable to colleges for institutional planning and financial stability, as well as to help reach their preferred financial, demographic, and scholastic outcomes for the incoming student body.

Unfortunately, the widespread use of enrollment management algorithms may also be hurting students, especially due to their narrow focus on enrollment. The prevailing evidence suggests that these algorithms generally reduce the amount of scholarship funding offered to students. Further, algorithms excel at identifying a student's exact willingness to pay, meaning they may drive enrollment while also reducing students' chances to persist and graduate. The use of this two-step process also opens many subtle channels for algorithmic discrimination to perpetuate unfair financial aid practices. Higher education is already suffering from low graduation rates, high student debt, and stagnant inequality for racial minorities—crises that enrollment algorithms may be making worse.

This paper offers a range of recommendations to ameliorate the risks of enrollment management algorithms in higher education. Categorically, colleges should not use predicted likelihood to enroll in either the admissions process or in awarding need-based aid—these determinations should only be made based on the applicant’s merit and financial circumstances, respectively. When colleges do use algorithms to distribute scholarships, they should proceed cautiously and document their data, processes, and goals. Colleges should also examine how scholarship changes affect students’ likelihood to graduate, or whether they may deepen inequities between student populations. Colleges should also ensure an active role for humans in these processes, such as exclusively using people to evaluate application quality and hiring internal data scientists who can challenge algorithmic specifications. State policymakers should consider the expanding role of these algorithms too, and should try to create more transparency about their use in public institutions. More broadly, policymakers should consider enrollment management algorithms as a concerning symptom of pre-existing trends towards higher tuition, more debt, and reduced accessibility in higher education.

---

## Introduction

Algorithms have played a role in college enrollment as far back as their use at St. George’s Hospital Medical School in the 1970s—with an algorithm that was later discovered to be discriminating against women and racial minorities. Despite this troubling omen, algorithms have grown continually more important to how colleges shape their incoming student cohorts, a process called enrollment management. According to a 2015 Educause Survey, over 75% of colleges and universities use analytics for enrollment management, up from just over 60% in 2012, making it the most common form of data analytics in higher education. Algorithmic enrollment management is primarily done through vendors, including EAB (serving around 150 institutions), Ruffalo Noel Levitz (or RNL, serving around 300 institutions), Rapid Insight (around 150 institutions), Othot (which lists 30 institutions), Capture Higher Ed (around 100 institutions), Whiteboard Higher Education, and others, although some colleges develop their own algorithms in-house.<sup>[1]</sup> This proliferation of algorithms to at least 700 institutions is not inherently problematic, as colleges have legitimate need to predict the number of students who will attend in a coming year, as well as to budget and prepare accordingly.

---

**“The algorithmic enrollment optimization process warrants additional scrutiny, especially since it may contribute to pre-existing crises in higher education.”**

---

However, there is cause for concern about using algorithms to determine scholarship offers for college applicants, an increasingly common practice. These algorithms help assign scholarships to maximize either net tuition or yield—the percent of accepted applicants who end up attending that specific college. Through a two-part process—first prediction, and then optimization—a college may compare from a handful to thousands of different scholarship disbursement strategies to reach their preferred financial, demographic, and scholastic outcomes for the incoming student body. While these algorithms tend to be effective in increasing net tuition and yield, the most profitable scholarship strategy may not be that which is best for student success. The algorithmic enrollment optimization process warrants additional scrutiny, especially since it may contribute to pre-existing crises in higher education, such as an increase in student debt burdens, higher dropout rates, and the failure of many colleges to proportionately enroll students of color.

## **Why are colleges turning to enrollment algorithms?**

Understanding the broader state of higher education helps illuminate why colleges are turning to algorithms for enrollment management. The yield rate at an average college fell nearly 15 percentage points to 33.7% from 2007 to 2017. This is partially driven by high school graduates applying to more colleges and has led to more furious competition over yielding applicants, including more active recruitment activities. Recruiting one undergraduate student at a four-year private college now has a median cost of around \$2,100. As tuition has risen steadily, so too have tuition discounts, such as grants and scholarships. On average, tuition discounts have doubled from \$10,000 to \$20,000 between 2008 and 2018 and now encompass 52.2% of institutional expenses. This suggests that colleges are more aggressively using both recruitment practices and price variation to entice enrollment.

This competition comes not just from private colleges—public colleges, especially selective, flagship, and land-grant institutions, also compete for the same students. Despite growth over the last eight years, educational support from state and local governments remains lower, on a per-student basis, than the high-water mark in the year 2000. To raise money and compete on educational offerings and amenities, public colleges have raised tuition, including a doubling of in-state tuition (before scholarships), and shifted the proportion of students towards those from out-of-state and other countries—all of whom pay more than in-state students. This data paints a picture in which the nation’s colleges are more active in using higher tuition and targeted scholarships to raise revenue.

## **How algorithms fit into enrollment management**

It is with this context in mind that one must consider the work done by college admissions and financial aid offices, and how it is affected by enrollment management algorithms.<sup>[2]</sup> College admissions and financial aid offices have critical financial goals weighing on their shoulders. These offices are solely responsible for accepting and yielding enough students to keep the college financially solvent. Only a tiny number of prestigious colleges have endowments or other income capable of supporting college functions—the vast majority are reliant on tuition. This creates enormous pressure on financial aid offices to yield enough students to cover institutional expenses without giving out too much scholarship funding, often while also juggling other goals, such as enticing a diverse and accomplished cohort of students. To do this, admissions offices first decide who to accept, possibly also assigning a categorization (e.g., strong accept vs. weak accept) or a number score about the quality of an application.<sup>[3]</sup>

Then, usually in a process separate from deciding who to admit, these offices allocate scholarships in an attempt to entice students to attend. Unsurprisingly, research shows that higher scholarships awards increase the likelihood of enrollment. However, it is challenging for financial aid offices to precisely allocate limited financial resources to achieve their goals, whether that be raising the most revenue, attracting the best students, or some combination of both. This is precisely where algorithms shine—they help financial aid offices make better educated guesses about how scholarship funding increases likelihood of student enrollment. In one academic study, this resulted in a 23.3% increase in enrollment yield (from 12.1% to 14.8%) of out-of-state students.<sup>[4]</sup> One major vendor advertises a “10% increase in enrollment” through scholarship optimization. This scale of impact could translate into millions of dollars

in additional tuition revenue, which is the most prominent reason that so many colleges are moving towards enrollment management algorithms. As an ancillary benefit, the use of these algorithms may also reduce the workload for financial aid and admissions offices facing many thousands of college applications, which in turn reduces time and labor costs.

## **The algorithmic enrollment management process**

Prior work from New America has discussed how predictive algorithms are used to estimate which applicants are likely to attend a college, the size of incoming classes, and the financial needs of those students. However, it is important to note that the *predictive* aspect is only half of a two-step process. There is also a second critical step, which is the *optimization* of decisions that the college can make, especially around providing scholarships to accepted applicants. While optimizing scholarship disbursement has historically been done manually, there is an emerging trend of using optimization algorithms, which raise additional concerns. Understanding this two-step process is a prerequisite to identifying the possible effects of enrollment management algorithms on higher education.

### ***Predicting likelihood to enroll***

The predictive step of algorithmic enrollment is aimed at estimating how likely an accepted applicant is to enroll in a specific college. To do this, a college will first consolidate data about a college's past applicants, including variables like their high school GPA, standardized tests scores, FAFSA data, how much financial aid they received, where they live, and demographic information. Colleges also frequently incorporate engagement metrics, such as how often applicants attend college recruitment events and what percentage of college emails they read. Using this historical admissions data, the college or a vendor will then build a predictive model with these variables to predict whether each of the accepted applicants chose to enroll. To be precise, this is an application of supervised machine learning, using methods that might include logistic regression (the most common cited by vendors), decisions trees, or neural networks. The resulting model is a complex set of associations between these input variables and whether past applicants enrolled in the college.

Once it is built, the model can be used to predict the likelihood of enrollment for future applicants—their likelihood of yield. For any new applicant, this predictive yield model can generate an estimated percentage likelihood of enrollment. The predictive yield model also

enables financial aid officers to adjust a hypothetical scholarship amount for any newly accepted applicant and see how that funding change affects their likelihood of enrolling. With the model, admissions officers might also determine which applicants are worth focusing more time and resources on (e.g., college branded apparel and glossy pamphlets). Notably, since applicants frequently apply to multiple colleges (80% apply to three or more and 36% apply to seven or more), or alternatively can opt to not attend college at all, these models are not particularly accurate for any individual applicant. However, when used to look across many applicants, it can provide highly valuable aggregate information, such as the total expected incoming students, or total expected net tuition. Because this model is valuable in the aggregate, it can also be used in the next step of the process—deciding on a strategy for distributing scholarships.

### ***Optimizing for financial, demographic, and scholastic outcomes***

Optimization is the second half of the enrollment management process, and it seeks to discover a strategy for engaging with applicants that will lead to the best enrollment outcomes as defined by that college. Because scholarship funding is the most effective lever, the most attention is typically given to how to allocate this aid to maximize applicant yield. To do this, a college needs the predictive model from the prior step, a dataset of newly accepted applicants to that college, and at least several potential scholarship disbursement strategies to compare; one scholarship disbursement strategy is a list of individualized scholarship amounts to be offered to each accepted student. A college will then compare a set of possible scholarship disbursement strategies and choose the one which is expected, according to the predictive model, to result in the preferred incoming cohort for the college. The process by which colleges choose potential disbursement strategies can vary, but there are two important categories to consider: (1) by-hand optimization and (2) algorithmic optimization.

#### By-hand optimization

At present, by-hand optimization is the more prevalent approach to exploring possible scholarship disbursement strategies. “By-hand” entails that individuals (as opposed to an algorithm) choose a scholarship strategy, evaluate its potential outcomes with the predictive yield model, and then iteratively make changes to the strategy and compare the results. For instance, the software systems created by the vendors EAB, RNL, and Rapid Insight enable financial aid offices to design a strategy for how much funding each applicant will get. This

strategy is then run through the predictive yield model to see how it is expected to affect enrollment. A strategy might be as simple as segmenting the students based on high school GPA and SAT scores, and then giving \$5,000 to the bottom third of applicants, \$10,000 to the middle third, and \$15,000 to the top third. Based on conversations with vendors and software demonstrations, an applicant's expected family financial contribution also plays a prominent role, in addition to their GPA and SAT scores. The proposed strategies can segment students by many other variables—vendors mentioned specific high schools, distance traveled to attend, Pell grant eligibility, and more. In emails and conversations, several vendors also described how their process can even allocate both a mix of need-based and merit-based aid.

Often, the vendors sell both software and consulting support to help colleges consider various scholarship disbursement strategies. The consultants and financial aid office work together, exploring various disbursement strategies and breaking down the resulting predictions. In an email, EAB notes that they often look at the effect of different strategies on “enrollment, diversity, academic profile, access or other considerations, along with net tuition revenue.” While the scholarship disbursement strategies are first chosen for the entire population of accepted applicants, individual adjustments can be made for any given student. This might reflect a special interest in enrolling individual students, based on their qualifications or college fit, or perhaps a result of negotiation by a student who has better financial aid offers from elsewhere. All the while, the financial aid office can see how much the predictive algorithm expects to dole out in scholarships and the characteristics of students it predicts will attend.

It is important to note that this approach can be time-consuming for consultants and admissions officers. They must choose some number strategies of scholarship disbursement, and then compare their predicted outcomes against one another. These employees cannot evaluate every conceivable scholarship strategy and must quickly (within days or weeks) make a decision about what scholarship strategy to implement—even if they may continue to adapt offers for individual students. In contrast, an algorithmic approach to optimization has no problem evaluating many thousands of possible scholarship disbursement strategies.

### Algorithmic optimization

Optimization algorithms are an emerging alternative to manually creating and comparing scholarship disbursement strategies. While it is not clear that this approach is widely in use, a [recent paper](#) by University of Washington (UW) researchers details how implementing this algorithmic step at a large unnamed public university improved out-of-state applicant yield by 23.3% (from 12.1% to 14.8%) over the pre-existing process. A [similar study](#) using data from the Southern Illinois University Carbondale (SIUC) simulated that this approach would raise yield substantially, although it was not implemented due to a leadership change at the university. Further, at least one enrollment algorithm vendor, Othot, has implemented this type of algorithmic optimization, for which the default setting is to minimize scholarship aid while maintaining enrollment. Other vendors are at least moving in this direction as they compete to offer new “prescriptive analytics,” which is the [industry’s term](#) for guiding marketing, outreach, and scholarship decisions based on predictive models. Algorithmic optimization is clearly effective and could be ubiquitous in the next generation of enrollment analytics, and thus warrants careful consideration.

Using optimization algorithms, as opposed to by-hand exploration, for this step of enrollment management requires four inputs: a dataset of new college applicants, the predictive model from the prior step, a list of constraints set by the college, and a single value to optimize (yield, in our case). Before the algorithm starts, colleges must decide on a set of constraints, which might include a total scholarship budget and a maximum scholarship for any individual student. These constraints will determine what scholarship disbursement strategies are eligible choices; a strategy that does not meet the constraints will be thrown out by the algorithm.

Next, the optimization algorithm begins its work. To start, it generates thousands of strategies for how to disperse scholarship funding across all applicants, subject to the college-set constraints. Then, within each option, the optimization algorithm uses the predictive algorithm to see how likely each applicant is to attend, given the scholarship they would hypothetically be offered. At this point, using the predicted probability of attendance for each applicant, the algorithm can then estimate the total number of expected students and total expected tuition for each scholarship disbursement option. These values—total students yielded and total tuition paid—are examples of what the algorithm is working to maximize. The algorithm may keep trying different options by combining strong performing strategies and continually comparing them against one another, before eventually deciding it has done its

best to find an optimal solution, based on its goal and constraints. This process is analogous to the by-hand alternative, except that because algorithms are tireless, they may evaluate thousands of potential strategies.

Since this process is automated, its outcome is highly dependent on the constraints selected and the value to be optimized. Thus, these are incredibly important choices by the college. Anything that can be translated into a number can be optimized for. Since outcomes like total yield and total net tuition are easy and aligned with the goals of admissions offices, these are the most likely choices. The constraints placed on the optimization algorithm are also critically important. For instance, the aforementioned papers from UW researchers and SIUC both used a constraint to require that students with higher application scores (as determined by the admissions and financial aid offices) would be guaranteed higher merit aid offers than students with lower application scores. Different constraints could impose demographic diversity, such as by race or gender, or even require that a certain mix of intended majors was expected to enroll—anything that can be quantified for every applicant can be used as a constraint. In addition to the process that generates the predictive model, the choices of what constraints to impose and what value to optimize for are hugely influential.

## **The potential harms of enrollment management algorithms**

Algorithms can play a responsible role in higher education enrollment management and are not inherently harmful. As just one example, predicting yield is valuable for important considerations like institutional financial stability, ensuring course availability, and preparing sufficient student housing. It is also important to note that algorithms are not fundamentally changing the incentives and actions of colleges, which engaged in enrollment management practices long before algorithms. Rather, algorithms are further empowering universities to be more precise and effective in enrollment management, especially when evaluating scholarships. Again, the outcome from this trend is not definitionally bad, as it is possible that codifying rules into algorithms can make scholarship awards more consistent and fairer, or could even be used to prioritize meeting student's full financial need. However, this is not automatically true, and it should not be assumed to be the case.

---

## **“These algorithms may be driving up the number of applicants attending college at the edge of their financial capacity.”**

---

Understanding the algorithmic process used in enrollment management helps clarify why their widespread use might be cause for concern, of which three issues supersede the rest. First, when these algorithms work as intended, they may reduce average per-student scholarship support. Second, it is problematic that these algorithms typically optimize scholarships for yielding students, rather than using scholarships to support student graduation and success. The third concern is the possibility of algorithmic bias, through which subgroups of applicants who appear to the algorithm to be less affected by changes in scholarship funding may be mistreated. While the data necessary to definitively confirm these problems is not public, the available evidence gives weight to these concerns. Given that so many American colleges are now applying enrollment management algorithms, their total impact is potentially dramatic, and thus they deserve careful consideration and further study.

### ***When the algorithms help colleges, they may hurt students***

The stated goal of enrollment optimization algorithms is to incentivize enrollment at the precise maximum tuition (or minimum scholarship) an applicant is willing to pay to attend that college. Vendors unanimously market their enrollment management software in this way—saying they intend to allocate the “minimum amount of aid necessary to meet and exceed your [enrollment] goals.” Therefore, the fact that these algorithms appear to be highly effective becomes a real concern, in that they may be driving up the number of applicants attending college at the edge of their financial willingness, and possibly capacity, to pay. It is important to note that this is not a guaranteed outcome—algorithmic enrollment could, on net, be convincing students to shift to colleges with higher aid offerings.<sup>[5]</sup> Yet this is not the most likely outcome, because the use of algorithmic modeling provides a significant informational advantage to colleges in negotiations with their accepted applicants and is therefore unlikely to systemically raise aid offers. One vendor case study from Othot claims that its analytics managed to enroll 173 additional freshmen at the New Jersey Institute of Technology, without a corresponding rise in scholarship aid. Another case study from EAB takes credit for a 33%

increase in net tuition and a six percentage point decrease in tuition discount rates (from around 22% to around 16%) at Aster University. This may be contributing to broader trends in higher education, especially the falling percentage of financial need met by colleges.

This effect will be especially pronounced when algorithms are used in the optimization step, since it compares thousands of scholarship strategies to find the most effective approaches. Notably, the implemented outcome of the UW study was a strategy of lower financial aid disbursements, which even led to the creation of a new lower scholarship tier (4-8% of tuition) that had previously not existed. The algorithms in the SIUC study similarly suggested reducing scholarships. Unfortunately for students, if algorithms succeed in their intended goal of effective scholarship allocation, they may also short-change students.

### ***Maximizing applicant yield, not student success***

Enrollment management algorithms are most frequently marketed towards improving yield—that is, getting a higher proportion of prospective students to attend a college. While other factors, such as scholastic quality, cohort diversity, and student retention are mentioned, the focus on tuition and yield is evident throughout vendor case studies, reports, and sales documents. The perspective of these vendors is critical, since the overwhelming majority of colleges procure algorithmic systems, rather than develop them internally. This viewpoint informs an algorithmic approach that is concerning, simply because of what it is not: optimizing for yield is a very different goal than optimizing for student success, retention, or graduation.

The scholarship that a student receives does not only affect a student's likelihood to enroll, it also affects their likelihood to graduate. One study found that \$1,000 in additional merit aid increased the graduation odds of a student by around 0.9%, with much larger effects for need-based aid. More dauntingly, an additional \$1,000 in unsubsidized loans *reduced* low-income students' likelihood to graduate by over 5%, which, the study notes, is “the largest negative factor for all aid estimates.” Beyond just the ability to pay, earlier research also suggests that scholarships can influence a student's attitude and level of commitment to college.

---

**“Given the high percentage of college students who drop out with debt they are unable to pay, it is a matter of national concern that hundreds of higher education institutions may be optimizing scholarships to entice students to attend, rather than succeed or even graduate.”**

---

Dropping out of college with student loan debt is the worst possible outcome of attending college, and yet is far too common. Only 62% of first-time full-time undergraduate students graduate from the institution they started at with a four-year degree within six years. A survey of 1,000 college dropouts found that, of those with loans, they held an average of \$14,000 in student loan debt, and just under half were making payments on that debt. Administrative data from the Department of Education estimates that half of non-completers have student debt, and puts their average amount of loans closer to \$16,000.<sup>[6]</sup> According to the Center for American Progress, 90% of people who default on student loans had received a Pell grant at some point, suggesting significant financial need in college. Clearly, undergraduate students are not getting all the scholarship aid that they need. Given the high percentage of college students who drop out with debt they are unable to pay, it is a matter of national concern that hundreds of higher education institutions may be optimizing scholarships to entice students to attend, rather than succeed or even graduate.

### ***Algorithmic discrimination***

Like many other algorithmic applications, such as algorithmic hiring or facial analysis, enrollment algorithms are susceptible to the possibility of biased outcomes—such as against racial minorities, women, people with disabilities, or other protected groups. Some vendors clearly encourage using SAT scores and GPA to help determine levels of scholarship funding, which may further wealth and racial disparities, although the inclusion of families’ ability to pay may have a countervailing influence. Students who only apply to a small number of colleges, and therefore have fewer choices, may also be less responsive to scholarships, leading to algorithms to deprioritize them for aid; although this could also be true for high-income

students who are similarly less responsive to scholarships. Other factors, especially prospective student engagement metrics, including their attendance at university events, could easily be correlated with demographic characteristics that would lead to bias.

Notably, the use of both predictive and optimization algorithms makes the danger of discrimination more complex than the case of purely predictive algorithms. In the two-step algorithmic process described in this paper, bias would come from a disparity in the effect of more scholarship funding on a student's likelihood to enroll. So, if providing more aid to Black applicants had a smaller estimated effect on their likelihood to enroll than providing more aid to white applicants, then one might expect biased outcomes against Black applicants from the optimization process. This might happen if, perversely, Black applicants needed more aid than white students to meaningfully raise their likelihood to enroll. Alternatively, applicants with certain disabilities may appear to an algorithm to be less engaged, and thus in turn less deserving of attention or scholarships. For instance, college web pages that are poorly designed for prospective students with vision impairments may result in those applicants appearing to an algorithm to be disinterested.

Without auditing specific college data and models, it is impossible to know if this will typically be the case for any specific protected classes or minority groups. However, the complexity of the algorithmic process, the many potential entry points for bias, and the separation between vendor-developed algorithms and college employees all contribute to the potential for discriminatory outcomes. This is especially concerning in light of significant existing racial disparities in higher education, even at public institutions. For example, [the Hechinger Report](#) identified fifteen public flagship universities that were underserving the state's Black population by at least 10 percentage points, and close to as many had the same issue with the state's Hispanic population. Black and Hispanic [students lag behind](#) white students in graduation rates too—by over 20 percentage points and 10 percentage points respectively. Colleges should take these algorithmic concerns seriously, not least because [civil rights law prohibits](#) discrimination in institutions that receive federal education funding.

### ***Other possible concerns***

A range of other concerns arise when considering the dramatic proliferation of enrollment management algorithms:

- *Metrics change behavior:* The metrics—values that colleges decide to count—attract attention. While colleges were clearly focused on yield and tuition before enrollment algorithms, building an entire algorithmic process that is primarily focused on maximizing yield is certain to change behavior. Infrastructure focused on yield is perhaps likely to detract from any focus on unique or hard to quantify applicant characteristics, creating a more standardized—and less creative or artistic—college application process.
- *Algorithmic drift:* The first year in which algorithms are used for enrollment management, they are based on historical data that predates the algorithm. However, in the second year, and even more so in each consecutive year, the data being used to build the algorithms has itself been derived from an algorithmic process. This means that future versions of the algorithms will tend to focus on where the college has been successful in the past, potentially creating blind spots and making the process resistant to changing circumstances or to new opportunities.
- *Vendor Herding:* There are a relatively small number (between five and 10) of prominent vendors in the enrollment management algorithm market, and on the whole, the advertised descriptions of their process and analytics are markedly similar. Since their processes seem relatively consistent, the outcomes might be as well—potentially leading to consistently good results for students who match the historical expectations of colleges, and consistently poor results for students who don't.
- *Human-algorithm interactions:* It is reasonable to wonder how well admissions and financial aid offices understand these algorithms. Because colleges do not develop them internally, it is not likely that many university employees have strong technical backgrounds in algorithms. In other applications of mixed algorithmic and human decision-making, algorithms have had unexpected and harmful impacts—and that potential certainly exists here. As just one possible example, a financial aid officer might overestimate an algorithm, believing that it is accounting for an applicant's challenging life or financial circumstances when it is not, leading that officer not to push for a higher aid award.

## Guidance and best practices for colleges

There are meaningful interventions that colleges can take to mitigate the potential dangers of algorithmic enrollment management. First, enrollment management algorithms should not be used in the admissions process, nor should they factor into need-based aid.

- *Admissions:* Enrollment management algorithms should not be considered at all during the admissions process. Many factors, including academic preparation, life accomplishments, cohort diversity, may factor into college admissions, but a statistical likelihood of enrolling should not be one of them, as this can only distract from whether a student deserves college admission on their merits.
- *Need-based aid:* Algorithms that allocate scholarships to drive yield should not be part of any process related to need-based financial aid—these goals are simply incompatible. Need-based decisions should be made solely on the basis of the prospective student’s financial necessity, taking into account how to best enable their collegiate success.

When using enrollment management algorithms outside of the bounds stated above, colleges should still exercise caution, as their implementation can have significant effects on individual students and the make-up of entire cohorts. The recommendations below offer a guide to ensuring that the worst harms of enrollment management algorithms do not come to pass.

- *Document algorithmic processes:* Colleges should comprehensively document the goals, principles, processes, data, and algorithms used in algorithmic enrollment management. The intended outcomes should be explicitly stated in advance of algorithmic development, and the implementation of an algorithmic process should clearly reflect the project’s goals and principles. This documentation should be available to, and understood by, all the college employees engaged in the financial aid process. Routine retrospectives should consider evaluating the algorithm using the college’s stated goals.
- *Examine historical data:* Colleges must carefully evaluate and consider the historical enrollment datasets on which they build algorithms. These algorithms are dependent on each individual college’s data, which means a college’s history may define its future. Colleges should carefully examine their historical input data to ensure that the correlations it contains are ones that college wants to perpetuate.
- *Evaluate student outcomes:* Colleges should require from vendors, or perform themselves, analyses that estimate student success (such as retention and graduation), under the scholarship disbursement strategies designed by the yield optimization algorithms. Many of the enrollment management vendors do also offer student success analytics, and some firms, such as Civitas Learning, are exclusively focused on student success. These checks can help colleges recognize if their more financially efficient strategies may also drive down student persistence and graduation. An even better solution would be for colleges to

rethink their optimization strategies entirely, and work towards an algorithmic approach that maximizes a combination of yield and student success metrics.

- *Use interpretable models:* Colleges should opt for more interpretable types of predictive models, such as logistic regression and decisions trees, over black-box models, such as neural networks. These models are far easier to examine, both for their broad function and their impact on individual students. This makes it easier for colleges to diagnose and, if necessary, change an algorithmic decision or revisit the algorithm itself. Interpretable models also tend to be less complex, which makes unexpected and unfair algorithmic behavior far less likely.
- *Perform bias audits:* Colleges should be actively introspective of their own algorithms, or alternatively be proactive in demanding bias audits from their algorithmic vendors. The challenge of algorithmic discrimination in predictive analytics is becoming better understood, and many tools exist for examining outcomes for bias. Public reporting of high-level results from these bias audits would also further transparency and help ease concerns from parents and applicants.
- *Human-led merit evaluation:* When ranking or segmenting prospective students for allocating merit-aid, colleges should not overly rely on metrics such as SAT scores and high school GPA. Using human scoring of application quality can account for a far wider variety of factors, such as volunteer service, extracurricular accomplishments, work experience, and resilience in the face of adversity. Human scoring would therefore enable an algorithmic process to reward accomplished, if not traditionally academically successful, applicants receive a fair amount of financial support. This would also mitigate fears of vendor herding, especially for unconventional college applicants. Colleges should also consider randomly selecting some scholarship awards to be determined with only human judgement, and then comparing the human recommendations to that of the algorithmic process, possibly preventing unreasonable algorithmic decisions.
- *Use constraints intentionally:* When using algorithmic optimization, colleges should consider using constraints to maintain student diversity—that is, requiring funding and expected yield to be well-apportioned across different groups of prospective students. For example, the constraints of the optimization process might be used to reject scholarship disbursement strategies that the predictive model suggest would lead to unbalanced populations—such as underrepresenting local students, racial minorities, or people of different sexes and gender identities.

- *Hire internal data science talent:* Colleges should not exclusively rely on vendors to understand enrollment management algorithms, and thus should have internal expertise that can explore and challenge vendor claims. By hiring at least one data scientist, colleges can better ensure their use of algorithms represents their broader institutional goals and the needs of their student body, rather than the metrics that vendors choose to communicate.
- *Maintain student financial support:* Finally, and perhaps most critically, in using algorithmic enrollment management, colleges should seek to become more efficient and fairer in their scholarship distribution, but not reduce overall per-student scholarships. If algorithms do lead to reduced scholarships, colleges should consider using the additional tuition to invest in other student support services, which are associated with higher student graduation rates.

## **Recommendations for policymakers**

Likely over 700 American institutions of higher education procure enrollment management algorithms, and others have internally developed them. The proliferation of these algorithms will likely continue, as holdout colleges may be disadvantaged in competition for enrollment. Unfortunately, there is very little public information about how these proprietary systems are designed. The research that led to this report entailed extensive engagement with vendor documentation and software trainings, as well as interviews with current and former algorithm developers, just to construct this relatively high-level description. Critical details are largely obscured by vendors and colleges, including descriptions of datasets, model types and parameters, evaluation processes, employee guidance, and student outcomes. Despite these notable uncertainties, the available evidence suggests that enrollment management algorithms are meaningfully affecting the higher education market, and are more likely to reduce, than increase, scholarship offers. In turn, it is possible that these algorithms contribute to colleges meeting less student financial need, higher debt burdens, student dropout, and racial disparities. These stakes help explain why the European Commission considers enrollment algorithms to be “high-risk” under its proposed artificial intelligence regulation.

---

**“[I]t is possible that these algorithms contribute to colleges meeting less student financial need, higher debt burdens, student dropout, and racial disparities.”**

---

Beyond the already mentioned steps that colleges can take, policymakers should demand more transparency in how enrollment management algorithms are used. State lawmakers can request public colleges and universities document the use and effect of these algorithms, and press vendors to provide more technical details. If enrollment and financial trends at public colleges are concerning—such as a falling percentage of need-met or stagnant racial disparities—lawmakers should consider commissioning an independent, third-party assessment of enrollment management practices, including the role of algorithms.

The Department of Education (ED) has a role to play, too, and should comply with recent OMB guidance to document what algorithmic systems might be covered by its enforcement and oversight capacities. In the future, ED could also survey colleges, or at least those that receive federal funds, to determine the scope of use of enrollment management algorithms. Based on this survey, ED could issue guidance on best practices for algorithmic enrollment management, especially regarding student success and student body diversity. More broadly, the availability of college financial aid data for policy researchers is seriously lacking, preventing much important analysis. Congress should take steps to enable and encourage more exchanges of student financial data for policy research purposes. This would shed light on the effects of algorithmic enrollment, as well as many other issues in higher education finance.

On a broader level, the wide use of enrollment management algorithms is symptomatic of pre-existing trends in higher education—towards greater competition for student tuition dollars. Perhaps policymakers should react not to the use of algorithms directly, but instead to the incentives that are driving colleges towards elaborate price discrimination to maximize tuition revenue. By better funding public higher education institutions and mandating that they lower tuition and better support the financial needs of local students, policymakers can improve college access and affordability. In turn, this will create downward pressure on the cost of

tuition at private universities, reversing recent trends. While the role of enrollment management algorithms warrants greater transparency and oversight, policymakers should not lose sight of the broader incentives that have led colleges to this point.

---

*The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.*

*Microsoft provides support to The Brookings Institution's Artificial Intelligence and Emerging Technology (AIET) Initiative. The findings, interpretations, and conclusions in this report are not influenced by any donation. Brookings recognizes that the value it provides is in its absolute commitment to quality, independence, and impact. Activities supported by its donors reflect this commitment.*

Report Produced by **Center for Technology Innovation**

## Footnotes

1. 1 Estimates from Rapid Insight and Capture Higher Ed provided over email.
2. 2 These may be one consolidated office or two separate offices within a university. This paper will use ‘financial aid offices’, with the intent of referring to the employees in charge of allocating scholarships for enrollment management.
3. 3 In some cases, algorithms are also used to help decide which students to accept. For instance, the company Kira Talent claims to be able to determine personality traits and fitness for colleges—a claim that should be treated with extreme skepticism, as algorithms are unable to effectively measure personality or emotional characteristics.
4. 4 In fact, the researchers were improving on an existing algorithmic enrollment approach implemented by a consulting company, that itself would (presumably) outperform no modeling at all. This is therefore possibly an underestimation of the ability to use algorithms to improve student yield.
5. 5 There is some noteworthy complexity here. If the enrollment algorithms raise yield with a set amount of scholarship funding, then definitionally there will be less aid on average per student. However, it is also possible that only the average per-student aid must be held constant, which may not lead to this same problem. Further, different financial aid strategies may entice applicants from other colleges, where they may have been offered a lower scholarship, or from the alternative of not attending college. All of these outcomes are possible, yet on net, it is likely that a targeted optimization approach will more often (than a person) identify the applicant’s willingness to pay, leading to lower scholarships on net. This intuition is bolstered by the results of the research papers cited, as well as the claims and statistics released by algorithmic enrollment vendors.
6. 6 Author's calculations using the Beginning Postsecondary Students Longitudinal Study for students who left without return, average cumulative debt six years after enrolling (2017). Data retrieved from: <https://nces.ed.gov/datalab/index.aspx>.