Final Ascendium Grant Report: New Strategies to Support Career Entry for Community College Graduates: Augmenting Intensive Career Advising Services with a Novel Job Recommendation Algorithm and Machine Learning

Brian Heseung Kim, Benjamin Castleman, Yifeng Song, & Alice Choe
July, 2022

Introduction

Individuals with college degrees earn substantially more than peers without degrees, especially in times of economic downturns such as the COVID-19 pandemic. The labor market and non-monetary benefits associated with earning a college degree have fueled decades of investment by the federal and state governments in various policies to improve degree attainment among lower-income and underrepresented populations: financial aid programs to increase college affordability; K-12 investments to strengthen academic readiness for college; and advising programs to support students and families to navigate the college and financial aid application process, among other policies and programmatic initiatives.

These investments notwithstanding, research on the intergenerational mobility of students attending college continues to show large and meaningful differences in the mid-career earnings of students coming from families in the bottom and top income quintiles (Chetty et al., 2017). Likewise, low socioeconomic status (SES) students receive substantially smaller returns to higher education than their high-SES peers, even conditional on earning a degree: High-SES students with a college degree see a 136% lifetime earnings premium over high-SES students with only high school diploma, whereas low-SES college graduates see a lifetime earnings premium of only 71% (Bartik & Hershbein, 2016).

Our own analyses show that lower- SES community college graduates in Virginia are substantially less likely to be employed at a living wage and earn less on average than higher-SES graduates from the same program of study and performance level (e.g., comparing high- and low-SES students who both earn a 3.5 cumulative GPA in an information technology associate's degree). For instance, among students in the top GPA quintile in the same college and the same program of study, Pell recipients earn roughly $2,200 less per year three years after graduation (Kim et al., 2022). In fact, when considering all programs besides nursing, Pell students in the top GPA quintile earn as much as non-Pell students in the bottom GPA quintile on average. These results suggest that even well-qualified low-income community college graduates may not have access to the same information, career advising, and/or social and professional networks as they navigate the career exploration and application process.

These disparities in economic well-being among students completing college and even graduating from the same college with the same GPA raise a fundamental question: Are investments to increase degree attainment among lower-income students sufficient to narrow longer-run economic inequality, or are investments to ameliorate the barriers that graduates encounter after college also necessary to ensure positive labor market outcomes and upward economic mobility?

The goal of this project was two-fold: (1) To develop an algorithm that generates highly personalized information about job matches for community college graduates as they enter the labor market; and (2) To design, implement, and rigorously test an intrusive career advising intervention that delivers those job matches to low-income community college graduates and supports them as they apply for jobs for which they are likely to be competitive and which would provide stable employment and compensation. Through this work, we hoped to address informational barriers and frictions that may differentially prevent low-income community college graduates from finding success and stability in the post-graduation labor market.

Section 1: Developing the Job Matching Algorithm

How the Algorithm was Constructed

The objective of the algorithm was to identify, for individual community college graduates, jobs that align with their program of study, are reasonably close to where they live, are posted close to when they would graduate, and that are posted by employers that have historically paid higher wages to community college graduates.

Data sources

To start, we focused on building the algorithm in the Virginia community college context. Specifically, the algorithm integrated four primary types of data: (1) student-level academic and demographic data, (2) historic employment and earnings data, (3) job postings data, and (4) occupation-specific average earnings data. We elaborate on each of these below.

1. **Student-level academic and demographic data**
   Source: Virginia Community College System

   These data included individual-level measures like program of study, academic performance (i.e., GPA), financial background (i.e., whether they were a Pell Grant recipient), and demographics (i.e., race and ethnicity). This information allowed us to identify jobs that are relevant to the training and credentials that students receive and to check that the algorithm does not produce biased job recommendations for different student groups.

2. **Employment and earnings data of all Virginia community college students**
   Source: Virginia Community College System (VCCS)

   These data helped identify employers that have historically provided strong employment and compensation to graduates from a particular program of study. Specifically, using data from the Virginia Community College System (VCCS), we used student-level employment and earnings data for years 2011-2019 to calculate the average compensation offered by firms to VCCS graduates.

3. **Occupation-specific average earnings data**
   Source: Bureau of Labor Statistics

   These data helped us to identify jobs that correspond to occupational codes with higher average salaries in Virginia.

4. **Job postings data**
   Source: Burning Glass Technologies

   We observed 5.1 million job postings in Virginia between 2011-2019. Of these jobs, approximately half (2.6 million) are appropriate at the applied associate's degree level or below – i.e., they do not require a bachelor's degree or above. Of those 2.6 million jobs, approximately 1.5 million align with a specific VCCS field of study. Among the approximately 1 million jobs remaining, some do not list a specific field of study required (or preferred) in their posting; others require (or prefer) a field of study not offered at VCCS.

   Only a third of the jobs that aligned with a specific VCCS field of study (approximately 500,000) matched to an employer for whom we can estimate average historic compensation. The remaining ~1 million jobs for which we could not estimate average historic compensation may be at employers that had previously not employed VCCS graduates; amplifying these positions could have potentially expanded the employment opportunities that VCCS graduates pursue.

Identifying and Ranking Jobs

Our intent was to identify currently available jobs for community college graduates that aligned with their program study, were close to where they live, and which offered stable employment and sufficient compensation. Accordingly, one design question that guided the development of the algorithm was how to weigh different measures of quality and relevance for different jobs.

Our approach was to construct multiple measures of quality and relevance at many different "levels" of specificity and aggregation. We used 19 measures of job quality and 18 measures of job relevance. Examples of measures of quality included average wages an employer paid to graduates of the same college (more granular) and average yearly compensation for employees in the same occupational code in the same MSA area (less granular). Relevance measures were each coded as binary, such as whether the job was posted in the same county as the student's ZIP code and whether the job was posted two quarters before vs. after the student's anticipated graduation quarter. Each member of our team independently assigned weights for each measure (i.e., a measure with a weight of 1 counts twice as much as a measure with a weight of 0.5). Though infrequent, any differences in assigned weights between team members were harmonized and reconciled. The resulting list of all measures and assigned weights can be found here.

We subsequently combined quality and relevance measures to sort (or rank) jobs. We provide a simplified example below with just three measures: listed salary, employer earnings quality, and average occupation earnings in Virginia. Given these three measures and their respective weights, Job A with a score of 0.16 is ranked above Job B with a score of 0.075.

	Listed Salary (W=0.8)	Employer Earnings Quality (W=0.75)	Average Occupation Earnings in VA (W=1.0)
Job A	0.7	???	-0.4
Job B	???	0.1	???

$$\text{Score}_A = (0.7 * 0.8) + (0 * 0.75) + (-0.4 * 1.0) = \mathbf{0.16}$$

$$\text{Score}_B = (0 * 0.8) + (0.1 * 0.75) + (0 * 1.0) = \mathbf{0.075}$$

Replication of job matching algorithm in Tennessee

We worked with our colleagues at the Tennessee Board of Regents (TBR) to replicate the job matching algorithm in Tennessee, with the goal of comparing how similar the algorithm performed in different state contexts. In most regards, the TBR team approached the algorithm development following the same approach we used in Virginia, e.g., they used the same structure and content of Burning Glass data and the same approach to ranking jobs on quality and relevance.

One difference is that the TBR team was not able to secure access to historic employment and earnings data from the state unemployment insurance system, so was not able to incorporate this data into the algorithm development. We view this as a useful learning opportunity about the cross-state consistency of the algorithm, since not all states would necessarily be able to leverage historic employment and earnings data. As a concrete illustration of this learning opportunity, one interesting observation is that the algorithm in Tennessee places much greater weight on listed salary in ranking jobs than does the Virginia algorithm. This makes intuitive sense, since in Virginia the job ranks are *also* driven by employer quality, which is not necessarily perfectly correlated with occupation-specific wages. One consequence of the greater reliance on listed salary in the Tennessee context is that the algorithm is more likely to trade off the alignment between students' program of study and a posted job in favor of a higher listed salary. This trade-off could be adjusted by fine-tuning the algorithm in Tennessee, so we do not view it as a major flaw of the algorithm replication. Rather, we believe this highlights ways in which the algorithm may perform modestly differently across contexts depending on the data sources available in each state.

Strengths and Limitations of the Algorithm

Strengths

We rigorously tested the algorithm along several performance metrics in data-driven simulations, using a historical sample of 662 students who graduated in 2017 from Tidewater Community College, Piedmont Virginia Community College, and Wytheville Community

Colleges. We highlight a couple of the main results here; a more detailed report about the performance of the algorithm can be found here. In addition, our colleagues in Tennessee generated the same performance metrics. In the Tennessee context, these performance metrics were based on 1,580 students who graduated from 13 community colleges in Spring 2019.

- **There was a sufficient number of jobs per student and per quarter to justify ranking them algorithmically.** In Virginia the test sample included a range of institutions that were situated in urban areas (e.g., Tidewater Community College) and also rural areas (e.g., Wytheville Community College). For each institution, the algorithm generated a sufficient number of jobs per student and per quarter - e.g., average of 769 jobs at Tidewater and average of 92 jobs at Wytheville. We observe a similar pattern in Tennessee: on average, the algorithm generated 508 jobs per student and quarter. The ample number of job recommendations justifies our decision to rank them algorithmically based on quality and relevance measures described above.

- **We did not observe that the algorithm produces biased recommendations to different student groups based on demographics or financial backgrounds**. It is possible that employer-level average compensation differs across dimensions such as students' race and ethnicity. For instance, the average compensation an employer pays to White students may be higher than the average compensation paid to Black students. We took steps to avoid coding these differences into the algorithm, since doing so could reinforce historic inequities by systematically encouraging Black near-graduates to apply for jobs that provide lower average compensation. Crucially, we did not observe any meaningful differences in the quality of job matches that the algorithm produces (e.g., employment stability, average compensation) along demographic and financial backgrounds like race, gender, ethnicity, whether they are a Pell recipient, or first-generation status, in either Virginia or Tennessee.

- **Jobs that the algorithm identifies as higher quality and more relevant correlated with metrics we would sensibly expect, such as higher salary and close to home.** Concretely, jobs that the algorithm ranks as more highly recommended fall into occupational categories whose median wage is substantially higher than jobs with lower recommended ratings. Similarly, jobs with higher recommendation ratings are more likely to be a perfect match with students' programs of study, temporally closer to student's graduation date, and geographically closer to students' home and college counties. This is largely true in both Virginia and Tennessee, though as we note above, given the Tennessee algorithm's greater reliance on listed salary, more highly-ranked jobs are occasionally a less-optimal match to students' program of study.

In summary, the algorithm performed as expected and without bias for the Virginia Community College System (VCCS) and for the Tennessee Board of Regents (TBR).

Limitations

We encountered a couple challenges during the development of the algorithm which we describe at a high level below.

Specificity vs. Precision and Coverage. One question we continually faced was how to balance specificity versus precision and coverage when calculating compensation at both the employer and occupation levels. Specifically at the occupation level, some occupations were common enough that the average earnings of these positions could be estimated as specifically as at the metropolitan statistical area (MSA) level – i.e., regions of grouped counties. This level of specificity allowed us to say something to the effect of, "People working as network administrators in the Charlottesville, VA metropolitan area earned average salaries of $87,140. We could then recommend these currently-open network administrator jobs near Charlottesville, VA that are a good match for your program of study."

In contrast, there were fewer common occupations for which average wages could only be estimated at only the state or even national level. The question then was, to what extent did we think the salary of a job posted in a given city/town in Virginia could be reasonably predicted by the Virginia average? More granular measurements might be better predictors, but we also reduced the precision of estimates (fewer and fewer individuals contributing to each estimate) and coverage (fewer and fewer occupations with sufficient sample) as we zoomed in on a given area.

A similar challenge arose when estimating an employer's average compensation for past community college students. Because our historical data did not include occupational information (e.g., what job a given student had at a given employer we observe them working with), we could only really look at the employers as a whole. Still, we tried to estimate an employer's average compensation *per program*. For instance, if we saw a Management-related job posted at Employer A, we might want to reference how well Employer A had paid Management students in the past, rather than just how well Employer A had paid all students in the past. Alternatively, we might want to zoom in even more and reference how well Employer A had paid past Management students *from the same college* as a current graduate. As mentioned above, however, greater specificity and relevance for employer compensation estimates commensurately reduced the precision and coverage of any ensuing estimates.

Reconciling Conflicting and Incomplete Data. Imagine you saw two jobs available for a given graduate; both did not include a listed salary – or, equivalently, both listed broad salary ranges that almost entirely overlapped. For one job, it was an occupation that typically pays well in Virginia. For the other job, the employer had historically paid other graduates of the same program well. In instances like this where the data was incomplete, it was unclear which one to prioritize for the graduate. We ultimately decided to address this issue in our algorithm design by creating expert-driven rankings for the relative importance of each datapoint when available, and these rankings then adjust the weight with which each datapoint factors into the final job indices. While creating these rankings was an inherently subjective process, we took several steps to ensure close agreement in the final rankings among the team of experts involved and went on to garner general approval from peer researchers we consulted.

Soliciting Feedback from Researchers and Practitioners

As part of developing the job matching algorithm, we actively sought technical feedback from data scientists, economists, and education researchers.

- **Iterative improvements**: Much of the feedback revolved around ways to fine-tune our existing approach. For example, an important decision we made in our current algorithm was how to think about the reliability of a job's quality measure when we had relatively few data points to infer its quality off of (e.g., if it was missing listed salary); our colleagues offered a few additional means of handling this uncertainty using other statistical methods. We also made as transparent as possible the fact that we put different weights on each available measure of job quality. Correspondingly, one suggestion we received was to adjust the weights based on the general availability of certain measures (e.g., reduce the weight of measures that are rarely observed).

- **Common sense checks**: Because recommendation engines can be obtuse and "black-box" in nature, it was important to find creative ways to check that the algorithm was operating as expected. We received a handful of suggestions centered on checking that certain job-level measures we utilized were statistically valid and reliable (e.g., the employer-level quality measures), and that the job recommendations the algorithm produced actually aligned with human judgment. The testing of the algorithm along performance metrics (described in the section above) was a direct result of this feedback.

- **Additional data sources and measures**: One of the main strengths of our algorithm approach was the ability to incorporate many disparate data points relating to job quality in our recommendations. As such, our colleagues were quick to offer guidance on other data sources and measures our algorithm could incorporate, such as "Commuting Zone" level workforce data (in addition to the Metropolitan Statistical Areas we already include), employment and workforce data specific to AA-level graduates (rather than whole-workforce data we currently include), and statistics related to occupation-specific and sector-level growth trajectories.

- **Implementation suggestions**: Our feedback sessions garnered lots of interest and excitement around how we might most effectively implement the algorithm in the field. For example, several colleagues suggested that we make the algorithm as "live" and flexible as possible, such that it could respond directly to student preferences for geography (i.e., regions to search within) and job types (e.g., part-time versus full-time). This feedback is reflected in the mock-up of the job matching interface we created and shared in a prior check-in call.

We moreover engaged over 50 career advising staff at the Tennessee Board of Regents (TBR) and the Virginia Community College System (VCCS) to gather feedback regarding the utility of the algorithm and strategies for getting students to engage with the job matches. We describe below the common themes that emerged from the three feedback sessions.

- **There are many career exploration and planning tools that colleges are using already,** including Handshake, Emsi, College Central Network, and Purple Briefcase. Many of the questions that advisors raised were around what makes this job matching algorithm different from existing tools. We also heard from one college that they developed their own campus-specific tool to support students with their career plans.

- **Staff see value in providing personalized career advising to students but do not have the capacity to do so.** For example, one of the community colleges has only 6 career advisors to serve 80,000 students. In addition, career advisors and staff handle multiple responsibilities that limit their ability to provide one-on-one guidance to students. However, one of the colleges shared that, since creating a dedicated career services program, both student engagement and advisors' capacity to provide meaningful support have increased.

- **The sorting and filtering features increase the utility of the job matching algorithm.** Advisors offered additional filter criteria based on their interactions with students, such as jobs or employers that are open to hiring people with disabilities, have a criminal history, and/or state-specific criteria (e.g., finding jobs that require a military security clearance is relevant to the Virginia context).

- **Leveraging faculty to help promote the job matches may be an effective approach to engaging students.** In addition to subject expertise and knowledge about jobs in their industry, faculty form deep relationships with students (e.g., during classes). Students may therefore be more responsive to encouragement from professors to use the job matching tool relative to a generic promotion from their college's career advising office. From an evaluation design perspective, however, randomizing at the faculty level would exacerbate the sample size concerns we describe in the next section.

- **Relative to providing personalized job matches to students, teaching psychosocial skills and knowledge may be more important to helping students navigate the job market successfully.** Students might not have the knowledge necessary for conducting a meaningful career and job search such as how to utilize job titles in their search or how to translate personal and academic interests into career pathways. Without these prerequisite skills and knowledge, students may not benefit from the job matching algorithm as much.

- **Related to the theme directly above, we should consider carefully when is the most appropriate time to introduce the job matching algorithm to students during their postsecondary journey.** Making the algorithm and job matches available to near-graduates might be too late. Introducing it earlier on (e.g., during the third semester) might be more effective in getting students to think proactively about their career options before graduating and navigating the job market.

Section 2: Deploying the Job Matching Algorithm

Designing an Intrusive Career Advising Intervention

In close collaboration with career advising experts within higher education systems and institutions, we aimed to design an intervention that combined intrusive career advising and nudges to deliver the algorithm's job recommendations to students and provide tailored support to lower-income student populations as they navigate the job market. Ideally, this intervention would have been co-developed with relevant partners to include the following components:

- A customized and dynamic web interface through which advisors could look up personalized job matches for individual students. This interface would have a separate student-facing view through which students could look up that same information on their own. The interface would update dynamically with user input - e.g., if a student changed the zip code, the list of job recommendations would update and re-sort accordingly. A draft mockup of a potential interface can be found here.

- An integrated communications plan that leveraged text, email, and each institution's college portal to encourage near-graduates to meet with a career advisor and to deliver information about personalized job matches to individual students - i.e., a list of the top 5 job matches for each student.

- Advising practices and strategies that are tailored to the unique assets of lower-income students and help mitigate barriers they face as they navigate the job market.

Within this design, we wanted to test two key variations of delivering information about personalized job matches to students: The first variation would push information about job matches directly to students - e.g., via text message or through their college portal. The second variation would have career advisors walk students through their job match information and provide ongoing advising support. Testing these two variations would have allowed us to explore human vs. algorithmic bias in career advising. Additionally, because intrusive career advising is more resource-intensive and thus has important implications for the scalability of the intervention, it was essential to evaluate whether that particular approach would yield differential impact than delivering information directly to students (without additional career advising support).

Key Considerations for Intervention Study Design

Small to moderate impacts

When thinking about the intervention design, a key determination was to project how large an impact it is likely to have. Importantly, several recent studies found that interventions that deliver personalized information, some in conjunction with lighter-touch support such as text message-based advising, produced small or even null effects on students' postsecondary outcomes (Avery et al., 2021; Bird et al., 2021; Gurantz et al., 2019a; Gurantz et al., 2019b; Sullivan et al., 2021). These recent findings suggest that this combined career advising and nudge intervention might similarly produce small to moderate impacts.

Minimum sample size

In order to inform the number of higher education institutions and near-graduates we would need to work with for the pilot study, we conducted a statistical analysis to determine the minimum sample size necessary to detect reasonably-sized impacts of a career advising and nudge intervention. In order to detect the small to moderate project impacts of the intervention with two treatment arms, we would need to recruit approximately 4,000 community college near-graduates from 8-10 colleges and over a 4-5 year period.

In a perfect world, we would have implemented the study design as described above. However, as with all rigorously evaluated interventions, we faced a number of risks and costs associated with implementation.

The Challenge: Implementation Costs and Fidelity Risks

Reaching the minimum sample size

As we noted above, we would need a minimum of 4,000 community college near-graduates from 8-10 colleges over 4-5 years. This level of recruitment is challenging given the pool of eligible near-graduates we could draw from is fairly limited. Specifically in the Virginia Community College System (VCCS), there would be approximately 4,000 non-nursing graduates from Associate of Applied Science (AAS) terminal degree programs.[1] Of these 4,000 non-nursing AAS graduates, approximately half (2,000) would be current or prior Pell recipients. There are moreover only a few institutions that graduate a larger share of AAS graduates, but then a long tail of remaining institutions where AAS students complete their degrees. Given this context at least in Virginia, we anticipated we would need to work with 8-10 colleges in the VCCS to reach 1,000 Pell-recipients in any given year. Correspondingly, we would need to add to the student sample every year for a minimum of 4-5 years in order to reach our desired overall sample size of 4,000.

Data coverage issues

Over the course of developing the job matching algorithm, we uncovered multiple data coverage issues that limit the quality of job information we can provide to graduates from some colleges and programs. For one, because many AAS programs have relatively few graduates in a given cohort, we had limited precision around providing information about employers' compensation histories that is tailored to graduates' respective colleges and programs. Additionally, many jobs in the Burning Glass data did not contain information on listed salary, further limiting the precision of information we could provide on the compensation that near-graduates could expect from a particular job posting. Ultimately, this meant there was a smaller share of near-graduates for whom the algorithm could identify high-quality jobs.

Prohibitively high costs

Effectively reaching near-graduates with personalized job match information and/or career advising services would require institution-level (rather than system-level) investments. For instance, each college would need to either pay for fractional shares of existing career advisors or hire part-time career advisors. Additionally, some institution-level engagement strategies–such as working with faculty to help promote the job matching interface–would require us to randomize students to one of the two treatment variations at a department or course level, rather than at the student level (i.e., it is unlikely that faculty would be willing to notify only certain students in their courses about the job match interface). More broadly, institutions' ability and willingness to support longer-term implementation of the intervention may vary over the years, including leadership, faculty involvement, career advising program structure, and budgetary flexibility. Such factors could escalate costs over the years.

Risk of implementation fidelity issues

---

[1] We focus on non-nursing programs in particular given that our analysis of socioeconomic disparities in labor market outcomes among community college graduates found no difference in outcomes among nursing graduates.

With multiple institutions and cohorts, the risk of intervention fidelity issues would grow considerably. For instance, the technological feasibility of embedding information about job matches personalized to individual students may vary from one college to another. In addition, the quality and commitment of career advisors may vary from one institution or cohort to the next; such variability could attenuate the efficacy of the career advising component of the intervention. Or, the baseline career advising services that institutions currently offer to near-graduates might differ, with some colleges already engaging in active outreach while others do not. Indeed, our direct experience with interventions in postsecondary education that have provided personalized information and advising (around community college transfer and college completion) have highlighted the challenges of ensuring fidelity to advising design across institutions and advisors (Bettinger et al., 2021; Bird et al., in progress).

Driving student engagement

A job matching interface and career advising program could potentially yield positive employment outcomes for low-income students only if it effectively drives them to engage with the intervention. Accordingly, one implementation hurdle that was particularly salient in this project was getting students to use the job matching tool and/or respond to invitations to engage with advisors. A growing number of studies highlighted the challenges of effectively engaging students through digital channels (Gurantz et al., 2019; Bettinger et al., 2021; Bird et al., in progress). Some further indicated that engagement is higher among those with higher baselines of motivation or academic performance, suggesting such virtual advising programs may be effective for more motivated or higher performing students but not for the highest need students (Gurantz et al., 2019; Avery et al., 2021). Even if we were to lower the barriers to engaging with the intervention (e.g., using an “opt-out” instead of an “opt-in” design), engagement levels would likely vary greatly. A recent study of student engagement with a text-based nudge intervention moreover revealed substantial variation in behavioral measures across students, such as the number of times students responded via text message and how long over the course of the intervention they remained engaged (Kim et al., in progress). These findings altogether underscore the importance of careful design in promoting the effectiveness of advising and nudging interventions.

Conclusion

In both Virginia and Tennessee, the job matching algorithm effectively identified personalized job recommendations for community college graduates that optimized quality and relevance considerations. College- and system-level career advisors saw value in these personalized recommendations but had concerns about their capacity to provide individualized advising to students. Combined with the concerns we highlight around sample size needs, implementation and evaluation cost, and potential intervention implementation fidelity issues, we do not advise proceeding with the development or implementation of an intrusive career advising intervention to deliver the job matching recommendations from the algorithm. That being said, we are currently collaborating with ideas42 to develop a proposal for the Heckscher Foundation for Children to identify existing and develop innovative and effective strategies for engaging students at broad-access colleges and universities, which could at least address the sample size challenge. More information about this developing proposal is available upon request.