



# What about the non-completers? The labor market returns to progress in community college☆☆☆



Matthew Zeidenberg<sup>a</sup>, Marc Scott<sup>b,1</sup>, Clive Belfield<sup>c,\*</sup>

<sup>a</sup> Abt Associates, 4550 Montgomery Avenue Suite 800, North Bethesda, MD 20814-3343, United States

<sup>b</sup> NYU Steinhardt - Humanities and Social Sciences in the Professions, 246 Greene Street, 3rd Floor, New York, NY 10003, United States

<sup>c</sup> Department of Economics, Powdermaker Hall, Queens College, 65-30 Kissena Boulevard, Flushing, NY 11367, United States

## ARTICLE INFO

### Article history:

Received 13 March 2015

Revised 15 September 2015

Accepted 16 September 2015

Available online 30 September 2015

### JEL:

I23

J3

### Keywords:

Returns to college

Non-completion

Postsecondary education

## ABSTRACT

Despite copious research on the labor market returns to college, very little has adequately modeled the pathways of non-completers or compared their outcomes with those of award-holders. In this paper, we present a novel method for linking non-completers with completers according to their program of study. We use this method to calculate the labor market returns to programs of study, accounting for those who obtain an award and those who do not. We use a large dataset of community college transcripts matched with earnings data. We find that different classification systems – by algorithm, intent or goal – yield very different enrollment patterns across programs. Importantly, these classifications make a substantial difference to earnings patterns. Returns vary by program completion and by program non-completion. Consequently, combining completers and non-completers yields a new pattern of returns. We find that the variance in returns by subject of study is reduced when we combine data on completers and non-completers. In particular, the large returns to nursing awards are substantially lower when we account for the probability of completing a nursing program and the returns to not completing a nursing program. In addition, progression *per se* does not lead to higher earnings for non-completers: progressing further in a nursing program is no different from accumulating general college credits. If validated, these findings have significant implications for policies on program choice and on student retention policies.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Copious attention has been paid to the labor market returns to particular education credentials (Altonji, Blom, & Meghir, 2012; Belfield & Bailey, 2011). But most students in two-year colleges and many in four-year institutions do not complete a degree or certificate program. For students who first begin in community college, only one-third will earn a credential from any institution within six years (with another one-fifth still enrolled). Moreover, these completion rates have been declining over time (Bound, Lovenheim, & Turner, 2010; Radford, Berkner, Wheelless, & Shepherd, 2010). Despite their being in the majority, little attention has been paid explicitly to these non-completers and their labor market outcomes. It is important to know how labor market

\* The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C110011 to Teachers College, Columbia University. We appreciate comments from Shanna Smith Jaggars. We are grateful the editing expertise of Amy Mazzariello and Doug Slater, both of which improved this paper immensely. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

\*\* For information about authors and CAPSEE, visit [capseecenter.org](http://capseecenter.org).

\* Corresponding author. Tel.: +1 718 997 5440; fax: +1 718 997 5466.

E-mail addresses: [mzeidenberg@gmail.com](mailto:mzeidenberg@gmail.com) (M. Zeidenberg), [marc.scott@nyu.edu](mailto:marc.scott@nyu.edu)

(M. Scott), [belfield@qc.edu](mailto:belfield@qc.edu) (C. Belfield).

<sup>1</sup> Tel.: +1 212 992 9407.

outcomes vary according to the students' status at point of exit from college: these students' status may vary with the number of credits they have accumulated, the courses taken, and their progression toward completing an award. Equally importantly, the outcomes of non-completers are essential for considering the returns to awards. Typically, estimates are reported only for the returns to those who actually completed a particular award and not those who intended to complete it. As such, these are *ex post* estimates of returns. But the optimal information for decision-making is the *ex ante* return to a college program, i.e. the return an enrolling student can expect. Only by combining evidence for non-completers with completers of a particular program is it possible to estimate the *ex ante* return. As well, non-completers vary in how far they have progressed in a program or even if they have enrolled in a program at all. For students who do not complete, it is unclear whether those who almost completed a specific program have higher returns than those who just accumulated college credits.

Identifying which programs non-completers are actually in is a challenge. Students fail to complete a college award for many reasons and many non-completers make very little progress toward completion. Broadly, there are two methods for categorization: stated preference, i.e. what the student proposes as their program of study; and revealed preference, i.e. what the student is actually doing to complete a program of study. Most attention has been paid to identifying students according to their declared program, with little research using revealed preference (Compton, Laanan, & Starobin, 2010). However, most of these studies have looked at completion rates and not labor market outcomes over the longer term.

In this paper, we look at labor market returns to completion and non-completion using a two-stage approach. Given the low average completion rate, and hence a greater and more varied group of non-completers, we focus on community college students for analysis. First, we utilize a novel method to determine a non-completer's program of study from her actual coursework. The method looks at the transcripts of all students who obtained each award in each program of study. It then classifies each non-completer with an award and subject corresponding to that of the set of completing students that each non-completer's transcript most closely resembles. In effect, the student is 'revealing' her program based on the courses she took in relation to courses needed to complete. Once each non-completer's program of study is identified, we apply a second method to determine how far along each student has progressed in their program of study. We compare this revealed preference approach with approaches based on stated preferences. In the second part of our investigation, we merge these data with data on labor market outcomes. This allows us to identify the labor market returns to non-completion pathways and to progress along each pathway. Also, by combining the sample of non-completers and completers, we are able to estimate the *ex ante* returns to awards and programs of study.

Our analysis is structured as follows. First, we describe the alternative approaches to categorizing non-completion and the challenges of categorization. We then set out our algorithmic approach. Next, we describe the data for analysis: a large-scale dataset of community college students with linked transcript and wage data over multiple cohorts and

spanning multiple years in the labor market. We report information on non-completers using the algorithm and then estimate labor market returns both for non-completers (by status and by progress) and in combination with completers to get *ex ante* results.

Briefly, our results are as follows. We find different enrollment patterns across the measures to capture student progress and these in turn affect completion rates. We find differences between *ex post* and *ex ante* returns and these vary across awards and subjects: awards and fields that are relatively lucrative remain so after adjustment for non-completion, but the differential shrinks. Finally, we find no evidence that progression matters: controlling for the number of credits, students who are further along in an award program have labor market outcomes that are no better than students who have just accumulated college credits.

## 2. Understanding college non-completion

### 2.1. Categorizing students who fail to complete community college

Conventionally, students have been categorized by award received, with the residual put into a single group called 'some college' or 'college dropouts'. Yet many – almost all – of these students intended or expected to complete their program. By 'program' we refer to two elements: the award the student is aiming for and the field of study (major); we focus first on awards and then on subjects. Programs vary in their completion rates. But they also vary in their progression rates: some students will drop out in the introductory classes; others will fail the upper level classes. Thus, it is necessary to identify which program a student is in even if (especially if) they do not complete that program; this is done using either stated or revealed preference methods.

The stated preference method identifies non-completing students according to their declared major or their proposed program (see, e.g., Compton, Laanan, & Starobin, 2010; Bailey, Alfonso, Scott, & Leinbach, 2004; Choy & Horn, 1992; Jacobson & Mohker, 2009; Stuart, 2009). However, not all students declare a major and often declarations are made late. For some non-completers a declared major might not correspond to their intended pathway but serves simply a place-holder, e.g. to satisfy college requirements. Even for completers, we have found that the last declared major does not always match their actual completed major: our data shows that in only 61% of cases does the initial major correspond to the final major. Other related approaches rely on students' declarations of intent or goal. However, these declarations may be inaccurate or missing: questions about intentions or goals are often presented in closed form with mixed options (e.g. the student might have to choose between 'associate degree' or 'transfer' responses); and sometimes the declarations have a default category for assignment. As Bailey, Jenkins, and Leinbach (2006) noted, it is often difficult to determine the intentions of a student. In fact, they may be unclear even to the student herself, especially when starting out. Generally, economists are skeptical about stated preferences, particularly in the context of an experience good with uninformed consumers (for information deficiencies of students, see Zafar, 2011).

Conventionally, revealed preference methods – looking at what students do – are preferable. The most common approach is to describe dropouts according to the number of credits they have obtained (Melguizo, Kienzl, & Alfonso, 2011). However, this approach assumes that all credits are created equal in how they relate to progress toward completion: no account is taken for repeated courses, or courses that cannot be applied to graduation, or general college requirements. Other studies have created taxonomies (e.g. Bahr, 2010). But these taxonomies are often restricted to a few categories and may not correspond to labels colleges use to distinguish students. Moreover, few studies have looked at labor market outcomes, with a focus on categorization of student pathways. In a recent and detailed investigation of the labor market using NLSY79, Agan (2013) separates out decision nodes based on college sector choices to yield eight different pathways. However, this evidence does not allow for non-completers to be linked to particular groups of completers such that *ex ante* returns can be calculated for programs. In most of these studies dropouts are identified *ex post* and, as a combined group, are separated from all completers.<sup>1</sup>

Our approach – emphasizing the revealed preference method – looks at student course-taking patterns and assigns non-completing students to a particular program based on awards and subjects of study. In other words, students are matching according to the amount of postsecondary human capital they have obtained. Using a classification algorithm, we can combine samples of non-completers with completers to calculate labor market returns *ex ante*; we can also look at how far students progress relative to completers.

## 2.2. The classification algorithm

In order to perform the classification, we use an algorithm commonly used in machine learning, the Naïve Bayes algorithm (Mitchell, 1997). This algorithm classifies entities with features into categories. In our case, the entities are the students as represented by their transcripts, the features are the individual courses that students take, and the categories are the awards that students earn. The job of the classifier is to take the transcripts of students who have not completed an award and predict what award they would have earned if they had continued on the pathway that the transcript implies. For the classifier, the input is a list of all features (courses) and whether or not that course is present or not; that is, a long vector of zeros and ones (a bit vector). Since there are many more unique courses in the community college system than any one student will have taken, the bit vector for any given student will be mostly zeros, with ones for the particular courses the student has taken.

The algorithm makes its predictions by compiling the probabilistic relationships between individual courses taken and the award earned, for those students who have earned an award. What makes the algorithm “naïve” is that it is based

on the assumption that the courses taken are independent of one another, that is, that the probability of taking a given course is independent of taking another one, or the probability of seeing them together is simply the product of their individual probabilities. In reality, this is far from the case; for instance, taking Nursing 101 greatly raises a student’s probability of following up with the upper-level Nursing 201 and Nursing 101 may even be a prerequisite for Nursing 201. But this dependency may not impair the predictions significantly. A student who has taken Nursing 201 is highly likely to be a Nursing major; information on their Nursing 101 enrollment is unlikely to change that determination (especially if this is a prerequisite for Nursing 201). In fact, machine learning researchers have found that the performance of the algorithm is quite good in many contexts, in spite of this naïve assumption, and therefore it is widely used (Rish, 2001). There are other learning algorithms that have been shown to have better performance (Caruana & Niculesu-Mizil, 2006), but the Naïve Bayes algorithm has the advantage of being relatively easy to use, easy to understand, and can be used as a “proof of concept” for the usefulness of classification of student transcripts.<sup>2</sup>

Formally, the algorithm can be expressed as follows:

$$Y \leftarrow \arg \max_k (Y = y_k) \prod_i P(X_i | Y = y_k)$$

Where the  $X_i$  are indicators for each feature (course), the  $y_k$  are the possible categories (awards). For each award, the algorithm computes the product of the probabilities of each course having the realized values, given that award in each case, and multiplies it by the overall probability of the award. The award  $Y$  corresponds to a particular value of  $k$ , where  $k$  is the subscript over which the awards range. The maximum value of this product is then selected by the algorithm.

To “train” the classifier, we used ten-fold cross validation (Kohavi, 1995). This means that we divided the training data (here, the completer’s transcripts represented as bit vectors) into ten equal-sized sets, randomly. We trained the algorithm ten times, each time with a different tenth of the data left out of the training. We then tested the performance of the algorithm on each corresponding, remaining tenth of the data. The prediction for each completer was the prediction for the run of the algorithm in which it was left out. This cross-validation guards against over-fitting to specific realizations of the data. From these data, the Naïve Bayes algorithm is able to create a prediction of what award appears most likely given every student’s transcript data, completer and non-completer alike.

We examine the validity of this revealed preference approach in several ways. For the completers, one test is the rate at which the prediction matches the actual award. It is not possible to assess its performance on the non-completers,

<sup>1</sup> We are not focused on demographic or college characteristics that are associated with non-completion. This literature is extensive (Desjardins, Ahlburg, and McCall, 2002; Doyle, 2009). This literature could be used to link completers to non-completers to estimate, for example, the *ex ante* returns to particular demographic groups. That said, we are not aware that this application has been extensively performed.

<sup>2</sup> In their highly-cited review of 10 algorithms, Wu et al. (2008, pp. 24–25) conclude: “This [NBC] method is important for several reasons. It is very easy to construct ... easy to interpret ... And finally, it often does surprisingly well: it may not be the best possible classifier in any particular application, but it can usually be relied upon to be robust and to do quite well.” See also Hand and Yu (2001). To check independence, we re-trained the algorithm using pairs of courses (e.g. Nursing 101 and 201) rather than singletons. This re-training did not materially affect our predictions (details available from the authors).

since we do not know what their “true” award would be, had they completed. For the non-completers, we compare the intended or final declared major (where available) and the predicted award. Of course, a gap between intent and our empirically-deduced major does not indicate that our method is incorrect. Also, for many students lacking a declared major, the algorithm is the only way we can discover any major information at all.

The closeness of the match is also important. This method can make either mild or severe errors of classification. For instance, if the method classifies a student in the AS in Business Administration instead into the Applied Arts and Sciences (AAS) in Business Administration, that is a mild error, because these two programs are very similar. On the other hand, if the student is classified instead into the AAS in Nursing or the AAS in Mechanical Engineering, those are much more severe errors. For the completers, we can get a sense of the relative number of mild and severe errors. It is difficult to do this for all of the misclassifications, because there are so many possible classifications. So we do it for some of the most common awards.<sup>3</sup>

There are several factors that might affect the predictive accuracy of the algorithm. First, many students do not progress very far in their studies. The fewer courses with which the algorithm has to work, the less precise it will be in assigning the student to a specific pathway. Only if a non-completer took only a course that no completer took does the model fail to predict a pathway. Of course, the fact that these students make little progress may suggest they had no specific pathway or intent other than ‘being in college’. Second, many programs have ‘back-loaded’ course requirements: for instance, most of the courses associated with a business program, such as accounting and economics, are taken in the second year of an Associate of Science program. Only for students who progress to take some of these required courses will the algorithm be accurate. Third, students may initially declare ‘generic’ or ‘default’ majors, such as liberal arts; they may plan to change their major when their interests get more specific but instead drop out. When colleges have many such declared majors, the model will be less precise. Fourth, many students take credits that are not required, either because they are unsure of the program requirements or because they cannot access the necessary courses (on average, completers take 11 more credit hours than necessary, see Romano, Losinger, & Millard, 2011, Table 5). This might introduce more noise into the data the algorithm uses. However, the lack of precision of the algorithm may be a consequence of students’ lack of specificity regarding their pathway, rather than a failure of the algorithm to identify specific pathways.

As well, we estimate the students’ progression along a particular program. Within their identified program, we

<sup>3</sup> The accuracy of the algorithm depends on the number of bins into which awards and majors are categorized; the more bins, the more possibilities for errors, all things being equal. Typically, colleges offer awards in the same field at different levels. One approach is to assign a student to one based on the relative proportions of the completer population: if most completers get AAS degrees, then most non-completers will be assigned to this award. One alternative is to collapse all awards into a given field and not be concerned about the specific award. However, there is a big difference in labor market outcomes between an AAS and a certificate.

define progression as the ratio of credits earned in the program’s course by a non-completer to the median credits earned by all completers of that program. Because completers do not all take the same courses (due largely to options and electives in each program), we use the heuristic of counting the credits earned only among the courses that account for the top 90% of enrollments in the selected program. Straightforwardly, the progression ratio tells us how far along the program the non-completer got as a percentage of what is required for completion.

We note two aspects of our method. First, our predictions do not necessarily reveal ‘preferences’ in the sense that students choose their programs. The predictions are based on what students actually do rather than what they want to do. This does not imply intentionality: students may simply take courses without much consideration as to what program they are in. But intentionality is not essential to revealed preference, which strictly speaking is ‘what happens’ rather than what students want to happen or think is happening. Second, our predictions can be discrepant from what the student is actually awarded. That is, the model may assign a student into a diploma program even if their award is actually a certificate. In this sense, the model is classifying awards in a different way to how colleges classify awards. It is classifying them based on the human capital they have accumulated through their course work rather than the college requirements for graduation.

### 2.3. Estimating labor market returns for non-completers

The evidence on the labor market returns to community college is now voluminous. In their summary, Belfield and Bailey (2011) report strongly positive returns across most awards: based on 18 studies, the average earnings premium for an associate degree compared with a high school diploma is 13% for men and 21% for women. Recent studies have estimated the returns to awards offered by community colleges using large-scale datasets (Jacobson, LaLonde, & Sullivan, 2005; Jepsen, Troske, and Coomes, 2014; Andrews, Li, & Lovenheim, 2012; Dadgar & Weiss, 2012; Liu, Belfield, & Trimble, 2014). These newer studies make comparisons within samples of postsecondary students (not between postsecondary students and high school graduates who never attended college). But the overall findings are similar: Returns to associate degrees are high, particularly in the health sciences, and gains for female students exceed those of male students. However, these studies reveal more heterogeneous returns to certificates and diplomas, as well as to variations across particular fields of study.

The studies that look explicitly at community college non-completers use an approach based on evidence of credits accumulated. A few studies have found earnings gains from credits or years of study that do not lead to a completed degree; gains are identifiable for as little as a semester’s worth of credits (Jacobson et al., 2005). As noted above, however, this approach does not provide a link between those who complete a program and those who do not. It also assumes that all credits are equal in terms of getting students through their program.

We estimate the returns to pathways using a standard Mincerian earnings equation:



$$\ln Y_t = \alpha + \theta AF_{t-k} + \beta X_{t-j} + \gamma Z_{t-k} + \delta EXP_t + \varphi EXP_t^2 + \varepsilon \quad (1)$$

Earnings  $Y$  are a function of: a vector of college awards/fields,  $AF$ ; a vector of pre-college characteristics,  $Z$ ; a vector of pre-college personal and ability-related characteristics,  $X$ ; and work experience,  $EXP$ . These variables are measured in quarter  $t$  and lagged to earlier quarters by  $k$  and are described in the Table Notes. We report quarterly earnings as percentage gains over baseline.

The coefficient  $\theta$  represents an estimate of the earnings premium from the program followed by the student in community college. For completers, the program is either an award (associate degree, certificate, or diploma) or award-with-subject (based on CIP categories). For non-completers, the program is identified using our classification algorithm. As a comparison, the program is identified using stated preferences of students about their award intentions and goals. Using this method we separate students into those who intended to get an award and did and those who intended to get an award but did not. We undertake an analogous separation based on student goals.

The coefficient  $\theta$  is an unbiased estimate of the gains from education programs if there is no omitted variable bias or selection/endogeneity bias (Brand & Xie, 2010). Although our earnings function does not fully address these biases, results using this method have proved to be highly robust and minimally biased (Carneiro, Heckman, & Vytlačil, 2011; Jepsen, Troske, & Coomes, 2014). As our sample all initially enrolled at a community college, we might expect less omitted variable bias than compared to using high school graduates as the baseline group (e.g. all subjects were motivated enough to enroll in college). Endogeneity bias may be present in our sample. All students faced similar labor market options because they began college at the same time (and in the same state) but clearly some entered the labor market before others and there may be unobservable reasons for this. Yet, it seems unlikely that the bias on unobservables is as large as the bias on observables (see Webber, 2014); and in earlier analysis we have found the bias on observables to be very small (Liu et al., 2014). Moreover, it is important to note that we are making comparisons between program completers and all program entrants and so the extent of the bias should be smaller.

Our estimation approach is as follows. First, to make equivalent comparisons, we identify as our program default category those students who obtained a diploma; all coefficients are therefore relative to these completers. Next, we estimate the returns to completers and the general class of non-completers, as per the general literature. We then estimate returns where non-completers are separated according to their program. Here the default category is individuals in diploma programs. This yields the returns to not completing a specific program.

We then combine the coefficients on completers and non-completers to estimate the *ex ante* returns. Specifically, we estimate these *ex ante* returns  $R$  as the weighted sum of the returns to completers  $C$  and non-completers  $N$ :

$$R_{ExAnte} = \varphi R_C + (1 - \varphi) R_N \quad (2)$$

Here,  $\varphi$  is the probability of completing the award or program. So, the *ex ante* returns to a pathway depend on the returns to completers, the returns to non-completers, and the probability of completing.

We expect that program estimates for the pooled group of non-completers with completers will yield lower rates of return to college. However, non-completers may have had more opportunity to work. The complementarity between education and experience may be such that over the window of analysis here the non-completers actually experience faster wage gains. Non-completers might also be inter-temporally substituting: taking employment when easily available with the expectation that they can return to college later if their earnings do not increase.

As well, we expect that the adjustment to *ex ante* returns will affect pathways differentially. Relatively lucrative award choices may not be so valuable once we adjust for differences in completion rates. Most likely, the adjustment should compress earnings gaps if high-earning programs are also those with the highest non-completion rates. For example, an Associate degree in Nursing leads to relatively high earnings; but if this degree program has a relatively high failure rate, then the net earnings advantage is attenuated.

Finally, we look at student progression to examined whether students who progress further along a given program have higher returns. We estimate a version of Eq. (1) for non-completers but include an indicator for progression  $PROG$ :

$$\ln Y_t = \alpha + \beta X_{t-j} + \gamma Z_{t-k} + \delta EXP_t + \varphi EXP_t^2 + \lambda PROG_t + \zeta CR_t + \varepsilon \quad (3)$$

Progression is almost certainly positively associated with earnings because students who have progressed further have more credits. Therefore, we include credits  $CR$  in our estimation. The coefficient  $\lambda$  is therefore identifying whether student progression toward an award yields higher earnings than simply accumulating college credits.

We expect that students who have progressed further on an award will have higher earnings than those whose progress was weaker, even after controlling for credits. Students who are progressing are most likely to have taken upper-level (harder) courses and mastered more material. They may also be more goal-directed such that the earnings gains identify unobserved motivation. Greater progression should lead to higher earnings. Alternatively, lower-level courses may produce more human capital with more general value in the labor market. Students who dropout might benefit from being generalists rather than specialists: taking courses in principles of economics and principles of sociology may be more useful than principles of economics and advanced economics, for example. Also, lower level and non-program courses are more readily available in the academic calendar; as such, students can more easily integrate their studies with their labor market activities.

### 3. Data

#### 3.1. Datasets

Our dataset is composed of all first-time-in-college, credit-seeking students in the North Carolina

Community College System (NCCCS) in the academic years 2002–2003 through 2004–05.<sup>4</sup> The NCCCS includes 58 colleges and enrolls approximately 100,000 new curriculum (award-seeking) students each year. The transcript dataset contains information on individual students, including full college transcripts (e.g., courses taken, grades earned, awards received, duration of study), basic personal information (e.g., age, gender, race/ethnicity), and financial aid received (loans and grants per semester). The transcript data do not include direct information on socioeconomic status or on students' prior academic achievement.

We merged the college transcript data with student-level data from the National Student Clearinghouse (NSC). The NSC tracks students as they transfer to other Title IV eligible colleges, as more than one third of all community college students do (Hossler et al., 2012). The NSC dataset includes information on awards subsequent to enrollment within the NCCCS but field of study is not reliably recorded across all colleges.

Our analysis focuses on the highest award obtained by each student at any institution within the designated time period. NCCCS awards include certificates, with 12–18 semester-hour credit requirements, which are intended to be completed in one year of full-time study or less; diplomas, with credit requirements ranging from 36–48 semester-hour credits, with a general education requirement; and associate degrees (mostly in applied science), with a required 64–76 semester-hour credits, including a cross-disciplinary general education requirement of 15 semester-hour credits. The NSC data allow for the identification of NCCCS students who eventually obtained a bachelor's degree or higher. These students are excluded from our sample.

Using Social Security numbers, we then merged this combined student dataset with earnings data obtained from the North Carolina Department of Commerce Unemployment Insurance (UI) records. The UI earnings data are collected on a quarterly basis from UI-covered employers and include total earnings from all jobs, as well as Standard Industrial Classification (SIC) and North American Industry Classification System (NAICS) information for each job (hours of work are not available). We use earnings data for the most recent available year (2011). Earnings are the average across quarters with non-zero earnings in 2011. All earnings are adjusted for inflation to be expressed in 2010 dollars using the quarterly Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W).<sup>5</sup>

<sup>4</sup> These data were used by Liu, Belfield, and Trimble (2015). We exclude students who were not enrolled in designated curriculum programs leading to awards but in customized training, personal enrichment courses, or other noncredit programs. For further details on the dataset, see Liu, Belfield, and Trimble (2015). Our models include first-semester GPA as a control variable. Omitting this variable does not materially affect our results or conclusions (details available from authors).

<sup>5</sup> Unlike sample survey data, UI earnings data are not affected by biases due to imputation, self-reporting, and nonresponse. However, the UI data do not include all workers; they exclude independent contractors, military personnel, some federal personnel, and those working in the informal sector (e.g., casual laborers). In most states, including North Carolina, state UI datasets do not include workers who moved out of state. However, in our sample over 90 percent had at least one UI wage record.

**Table 1**  
Descriptive statistics.

|   | Female      | Male        |
|---|-------------|-------------|
| <b>Awards</b>                             |             |             |
| Associate degree                          | .19         | .13         |
| Certificate                               | .04         | .07         |
| Diploma                                   | .03         | .02         |
| <b>Students predicted per award</b>       |             |             |
| Associate degree                          | .49         | .50         |
| Certificate                               | .31         | .34         |
| Diploma                                   | .20         | .16         |
| <b>Intentions</b>                         |             |             |
| Associate degree (intended and award)     | .04         | .01         |
| Associate degree (intended but no award)  | .13         | .05         |
| Certificate (intended and award)          | .04         | .07         |
| Certificate (intended but no award)       | .57         | .48         |
| Intended other                            | .32         | .42         |
| <b>Goals</b>                              |             |             |
| Associate degree (goal and award)         | .09         | .05         |
| Associate degree (goal but no award)      | .24         | .19         |
| Certificate (goal and award)              | .01         | .01         |
| Certificate (goal but no award)           | .04         | .03         |
| Goal other                                | .63         | .71         |
| Program match                             | .28         | .31         |
| <b>Credits earned</b>                     |             |             |
| Associate degree – award                  | 73.0 (25.3) | 68.3 (25.4) |
| Certificate – award                       | 35.0 (22.8) | 26.6 (20.0) |
| Diploma – award                           | 59.5 (20.4) | 53.4 (23.2) |
| Associate degree – predicted but no award | 25.2 (23.1) | 22.4 (22.2) |
| Certificate – predicted but no award      | 9.9 (12.9)  | 7.8 (10.9)  |
| Diploma – predicted but no award          | 18.0 (18.9) | 14.3 (17.0) |
| <b>Progression</b>                        |             |             |
| Associate degree – predicted but no award | 0.32 (0.29) | 0.29 (0.29) |
| Certificate – predicted but no award      | 0.30 (0.37) | 0.25 (0.34) |
| Diploma – predicted but no award          | 0.29 (0.30) | 0.25 (0.31) |
| Ln (Quarterly Earnings) <sup>a</sup>      | 8.16 (1.20) | 8.35 (1.24) |
| <i>Total observations</i>                 | 256,615     |             |

<sup>a</sup> Non-zero earnings observations only.

### 3.2. Descriptive frequencies

Table 1 shows the sample for analysis and information based on application of the algorithm (see also Appendix Fig. A1). Our sample includes students who receive an award at the community college level and then transfer to a four-year college and includes students who transfer but receive no award. We exclude students who transfer and then receive a BA degree without receiving a credential at the community college level. We assume these students are not intending to complete an award in community college level and so their progression status is not relevant.<sup>6</sup> As is the case across the community college sector, fewer than half of the students receive an award (excluding BA degrees). For Associate degrees, 19 [13%] of the female [male] students have received an award. Based on the algorithm, 49 [50%] of all students

<sup>6</sup> It is unlikely that we can make accurate predictions of awards and fields for students who do not intend to complete a community college award. For students who transfer but do not complete a BA we are less certain of their pathway. The rational pathway for these students – who actually terminate their postsecondary education with no award – would have been to complete a community college award: courses are cheaper and a community college award acts as insurance in case of non-completion. Empirically, we find that our results are very similar if we exclude students who transfer to a four-year college despite having no community college award (details available from authors).

were in Associate degree programs. So, the completion rate for Associate degrees is 39 [26]%. For certificates, 4 [7]% of students received an award. Yet 31 [34]% of students were in certificate programs. The completion rate for certificates is therefore 13 [21]%, considerably below that for Associate degrees. Finally, 3 [2]% of students completed diplomas, with 20 [16]% of students in such programs. The graduation rate is therefore 15 [13]%. Interestingly, these graduation rates are quite different from rates derived from IPEDS data (which do not count students who transfer) and from comparative analyses of colleges based on their proportions of awards that are certificates (Bailey and Belfield, 2012).

Two stated preference approaches are available in our dataset. One is a statement of the students' intentions; using a closed form question, students could declare their intention as an Associate degree or occupational award (which we interpret as a certificate), or other intentions. The other is based on the students' goals; also a closed form question, students could declare an Associate degree or skills upgrading (which we interpret as a certificate). As shown in Table 1, these stated preference metrics yield much lower rates of associate degree program enrollment and imprecise rates of certificate program enrollment. They also yield many uninformative 'other' codings; as shown in Appendix Fig. A1, a large proportion of students cannot be identified using these metrics. These factors suggest caution in using stated preferences as indicators of what students do. Given the ambiguity between intentions/goals and awards, we focus on the revealed preference measures for Associate degree students (and not students in other programs).

Finally, Table 1 shows the credits earned and progression rates for students by award. Credits earned are the typical indicator of student progress. For each award program, non-completers have about one-third of the credits of completers. However, this credit comparison may be misleading if the credits are not moving students toward completing a program. A more accurate guide is the progression rate. This is calculated as the ratio of credits earned in a program to the median credits earned by all completers of each program.<sup>7</sup> Looking at progression, non-completers do not get very far along their pathway: on average, non-completers only get about 30 [26]% along the way to completing their intended award.<sup>8</sup>

## 4. Results

### 4.1. Model accuracy

We begin by examining the model accuracy. Our first observation based on the computation of the algorithm is that there are very large numbers of awards and majors within the college system: we estimate there are over 600 combinations. In turn this suggests that students are unlikely to know *ex ante* which program to choose and, in their first year, which program they are making the most progress toward.

<sup>7</sup> As noted above, credits earned in a program are only those from courses accounting for the top 90 percent of enrollments in the program.

<sup>8</sup> Frequencies for the subsample with earnings data is given in Appendix Table 1. This subsample has very similar characteristics to the full sample on which the algorithm is applied.

Second, only 10% of students do not declare a major (the declaration rate is actually slightly lower for completers). This rate is sufficiently low that administrators may prefer to rely on the major for making program assignments. However, as noted above, the final declared major of completers only matches their final award in only six out of ten cases. So the major is not a very good guide to what awards the students end up with. It is likely to be even less useful a guide as to the pathways of non-completers. By contrast, the algorithm correctly classifies completers based on their courses 71% of the time.

Given these rates and measurement errors, the match between the algorithm and the declared major is low. The declared major matches the predicted major only 15% of the time. However, this low match rate is a function of both the bin numbers and the noise factors referred to above, as well as the trained target of the algorithm, which is the actual major. There are 637 different bins, many of which are very close to each other. The main factor reducing precision is that most of the non-completers have made very little progress (9 credits at the median).

The model does generate predictions that are, for some students, discrepant from their actual award. However, these discrepancies are few. Only 5% of students with Associate degrees were not predicted to receive this award; 10% of students who received a certificate were not predicted to receive one; and 5% of students who received a diploma were not predicted to receive one. In our analysis, we use our predictions of awards rather than the actual awards.

### 4.2. Labor market returns for non-completers

Table 2 shows coefficients for earnings gain relative to students whose highest award is a diploma from community college. (See Table Notes for covariate controls). The sample includes students who drop out of community college and students who transfer to another institution unless those transfer students then go on to complete a BA degree at a four-year institution (and so are better matched with four-year students). Model (1) aggregates all non-completers into a single group. The other three models differ in how they classify students *ex ante*: Model (2) uses the prediction from the algorithm; and Models (3) and (4) apply versions of students' proposed awards.

For female students, Model (1) shows the gains over diploma-holders for associate degree holders are +4% but for certificate holders they are -33%. Grouped together, the non-completers earn 32% less than diploma-holders. Therefore, we identify very high returns to *completing* an associate degree (+36%) or diploma (+32%) and insignificant returns to certificates (-1%). This last result is plausible: non-completers and certificate-holders have similar numbers of credits. These premiums correspond to the *ex post* premiums that are typically reported in rate of return studies.

Model (2) shows earnings gaps accounting for the predicted program of each student (again relative to diploma-holders). Students with associate degrees or certificates obtain approximately the same premium/deficit as in Model (1), with the exception of the Associate degrees for females. But the non-completers are disaggregated according to which award program they intended to complete based on

**Table 2**Returns to awards: actual, predicted, and stated preference  
Average quarterly earnings gains in 2011 for 2002–05 NCCCS Cohorts.

|                                  | Female               |                      |                     |                      | Male                |                      |                     |                     |
|----------------------------------|----------------------|----------------------|---------------------|----------------------|---------------------|----------------------|---------------------|---------------------|
|                                  | (1) Actual           | (2) Predicted        | (3) Intended        | (4) Goal             | (1) Actual          | (2) Predicted        | (3) Intended        | (4) Goal            |
| AA degree                        | 0.037**<br>[0.018]   | 0.088***<br>[0.016]  |                     |                      | 0.153***<br>[0.029] | 0.149***<br>[0.025]  |                     |                     |
| Certificate                      | -0.403***<br>[0.024] | -0.342***<br>[0.023] |                     |                      | -0.008<br>[0.032]   | -0.018<br>[0.027]    |                     |                     |
| Did not complete an award        | -0.386***<br>[0.017] |                      |                     |                      | -0.014<br>[0.028]   |                      |                     |                     |
| Predicted AA degree (no award)   |                      | -0.275***<br>[0.017] |                     |                      |                     | 0.008<br>[0.024]     |                     |                     |
| Predicted certificate (no award) |                      | -0.446***<br>[0.017] |                     |                      |                     | -0.049*<br>[0.025]   |                     |                     |
| Predicted diploma (no award)     |                      | -0.343***<br>[0.017] |                     |                      |                     | -0.070***<br>[0.026] |                     |                     |
| Proposed AA degree (awarded)     |                      |                      | 0.450***<br>[0.017] | 0.161***<br>[0.024]  |                     |                      | 0.108***<br>[0.041] | 0.148***<br>[0.033] |
| Proposed AA degree (no award)    |                      |                      | 0.014<br>[0.015]    | -0.239***<br>[0.023] |                     |                      | 0.002<br>[0.032]    | -0.013<br>[0.031]   |
| Proposed certificate (awarded)   |                      |                      | -0.031<br>[0.026]   | -0.192***<br>[0.068] |                     |                      | 0.055<br>[0.043]    | -0.052<br>[0.058]   |
| Proposed certificate (no award)  |                      |                      | 0.109***<br>[0.019] | -0.227***<br>[0.026] |                     |                      | 0.092**<br>[0.039]  | -0.012<br>[0.034]   |
| Proposed other                   |                      |                      | 0.092***<br>[0.020] | -0.164***<br>[0.023] |                     |                      | 0.087**<br>[0.040]  | 0.016<br>[0.031]    |
| Observations                     | 96,221               | 96,221               | 96,221              | 96,221               | 62,769              | 62,769               | 62,769              | 62,769              |
| R-squared                        | 0.135                | 0.137                | 0.121               | 0.123                | 0.165               | 0.166                | 0.163               | 0.164               |

Notes: Log average quarterly earnings in 2011 as dependent variable. Education coefficients are relative to student with diploma award. Cohorts of FTIC students from Fall 2002 to Summer 2005. Model includes: race (3), single parent, HS graduate, disability, enrollment age (2), experience and experience squared (measured as quarters with non-zero earnings age over 18), financial aid EFC, grants, loans, and aid, college GPA first semester, and college fixed effects. Robust standard errors in brackets.

\*\*\*  $p < 0.01$ ,\*\*  $p < 0.05$ ,\*  $p < 0.1$ .

the algorithm. All non-completers earn less, but the deficit varies according to program track. Non-completers on the Associate degree track earn 24% less than diploma holders, those on the certificate track earn 36% less, and those who do not complete their diploma earn 29% less. In other words, for students who terminate their postsecondary education with no award, those who are predicted to complete certificate programs have the worst labor market outcomes.

An alternative approach to estimating *ex ante* returns is to look at student proposals of intent or goal. As noted above, these proposals may be questioned: some students do not make them, make them late, or are assigned them by default; and they reflect students' stated preferences rather than revealed behaviors. Hence, we briefly review results for Associate degree students as shown in Models (3) and (4). For female students, there are very high returns to students who intend to get an Associate degree and actually do get one (+56% in model (3)). Analogously, Model (4) shows there are positive returns to students whose 'goal' is an Associate degree and that student actually gets an Associate degree (but not as strong, at +18%). Unsurprisingly, students who do not complete do considerably worse. Students who 'intend but fail' to get an Associate degree have earnings that are indistinguishable from diploma-holders (Model 3). Students who fail in their 'goal' of getting an Associate degree earn considerably less (-21%). Thus, the penalty for non-completion is

very different depending on whether intent or goal is used. There is a similar inconsistency between intent and goals in the returns to certificates, in part because these two indicators classify students at very different rates.<sup>9</sup>

The results for males are quite different than for female students. However, we caution that these differences may be because the returns to diplomas may be different across the genders. For males, *ex post* returns to Associate degrees are very high (+17%); returns to certificates and to the general group of non-completers are not statistically significantly different from diploma-holders. As shown in model (2), the returns to predicted non-completers in each award group are similar to those of diploma-holders. For students who do not complete, it does not appear to matter what they fail to complete. More starkly, the only pathway that does yield an earnings premium is completion of an Associate degree; for all other pathways (diplomas, certificates, non-completion) there is no statistically significant difference.

For the stated preference results, the premiums are very similar for students who either intend to, or have the goal of, getting an Associate degree and are successful (+15% and

<sup>9</sup> By comparison with our revealed preference approach, the outlier approach is 'intent': this last measure yields a zero penalty for non-completion relative to a diploma; the other two measures yield similar penalties (of 24% and 21%).



**Table 3**

Returns to awards: *Ex Ante* actual, predicted, and stated preference.  
Average quarterly earnings gains in 2011 for 2002–2005 NCCCS cohorts.

|                                      | Female               |                      |                     |                     | Male                |                     |                     |                     |
|--------------------------------------|----------------------|----------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                                      | (1) Actual           | (2) Predicted        | (3) Intended        | (4) Goal            | (1) Actual          | (2) Predicted       | (3) Intended        | (4) Goal            |
| <i>Relative to diploma enrollee:</i> |                      |                      |                     |                     |                     |                     |                     |                     |
| AA degree (awarded)                  | 0.337***<br>[0.011]  |                      |                     |                     | 0.203***<br>[0.016] |                     |                     |                     |
| Certificate (awarded)                | −0.101***<br>[0.020] |                      |                     |                     | 0.041*<br>[0.022]   |                     |                     |                     |
| Non-completer (excl. diploma)        | −0.079***<br>[0.010] |                      |                     |                     | 0.046***<br>[0.014] |                     |                     |                     |
| AA degree (awarded or not)           |                      | 0.132***<br>[0.010]  | 0.108***<br>[0.014] | 0.068***<br>[0.013] |                     | 0.104***<br>[0.014] | 0.042*<br>[0.023]   | 0.069***<br>[0.017] |
| Certificate (awarded or not)         |                      | −0.159***<br>[0.011] |                     |                     |                     | 0.016<br>[0.015]    |                     |                     |
| Other group                          |                      |                      | 0.026**<br>[0.011]  | 0.031***<br>[0.012] |                     |                     | 0.073***<br>[0.013] | 0.064***<br>[0.015] |
| Observations                         | 96,221               | 96,221               | 96,221              | 96,221              | 62,769              | 62,769              | 62,769              | 62,769              |
| R-squared                            | 0.132                | 0.124                | 0.115               | 0.115               | 0.165               | 0.164               | 0.164               | 0.163               |

Notes: Log average quarterly earnings in 2011 as dependent variable. Education coefficients are relative to student predicted to be on diploma track. Other group includes all other students except those on diploma track. Cohorts of FTIC students from Fall 2002 to Summer 2005. Model includes: race (3), single parent, HS graduate, disability, enrollment age (2), experience and experience squared (measured as quarters with non-zero earnings age over 18), financial aid EFC, grants, loans, and aid, college GPA first semester, and college fixed effects. Robust standard errors in brackets.

\*\*\*  $p < 0.01$ ,

\*\*  $p < 0.05$ ,

\*  $p < 0.1$ .

+18% respectively). This is despite these two declarations yielding very different proportions of completers (Table 1). As with the revealed preference approach, other pathways are not statistically different.

#### 4.3. Labor market returns ex ante

We now turn to calculations of the *ex ante* returns to each award. Using the specifications in Table 2 we can calculate the expected quarterly earnings across the combined group of both award completers and those predicted for that award. Based on the coefficients in Table 2, it is clear that the combined earnings are going to be much lower than those for completers.

To accurately identify the differences between *ex post* and *ex ante* premiums, we estimate the specification as per Table 2 but define the default group as all students who are predicted to be on a diploma track (not just those who complete one). We focus on the relative earnings for Associate degree programs and diploma programs. The results are given in Table 3.

The results in Model (1) show that, relative to students on a diploma track, students who earn Associate degrees have much higher earnings subsequently. For females, the gap is 41% and for males it is 23%. These premiums are much higher than in Table 2. This is unsurprising because the default group now includes students who failed to complete a diploma. Nevertheless, even those who complete a certificate have lower earnings than those who enroll in a diploma program (although the effect is not statistically significant for males).

Our primary comparison is between the actual results for Associate degree-holders and the evidence for those who are predicted, intend, or have the goal of being an Asso-

ciate degree recipient. These results are given in models (2), (3), and (4), respectively. Compared to those on the diploma track, the earnings gains for those predicted to be in an Associate degree program are still positive and statistically significant. They are 15% [12%] higher for female [male] students. This figure is much lower than for the actual completers (model (1)): *ex ante* returns are always lower than *ex post* returns. But critically this figure is much higher than the relevant model (1) comparison in Table 2. The *ex ante* return to an Associate degree is significantly positive and greater than the *ex post* return. We emphasize that this is all relative to a diploma. Looking only at students who complete Associate degrees and diplomas *under-states* the premium to the former. Adjusting for the probability of completing these respective awards, the Associate degree premium is *higher* (coefficients in models (1) of Table 2 and model (2) of Table 3).

An analogous result holds for certificates. For females, there is a large penalty for completing a certificate relative to completing a diploma (−29%, model (1) of Table 2). But this penalty is much lower when we adjust for who is in which program. The penalty for being in a certificate program relative to being in a diploma program is only −15%. That is, completing a certificate appears to be the worse type of award to complete but the disadvantage is not as bad when we account for the higher probability of completing one. This results holds for females; for males, the results are inconclusive as few pathways appear relatively lucrative.

These findings are partly corroborated when we use the proposed measures of intent and goal. When we estimate an equation where diploma enrollees are compared to those who intend to and do get an Associate degree, the returns to the degree are very large and comparable to those in column (1) of Table 3. The same is true for those whose goal is an Associate degree and they are successful in this goal. (These

**Table 4**  
Returns to awards: *ex post* and *ex ante*.

|                   | Returns relative to diplomas |       | Completion probability | Interpretation   |
|-------------------|------------------------------|-------|------------------------|--|
|                   | $R_C$                        | $R_N$ | $\beta$                |  |
| Female:           |                              |       |                        |  |
| Associate degrees | >                            | >     | >                      | Returns even <u>greater</u> after adjusting for non-completion                               |
| Certificates      | <                            | <     | <                      | Returns even <u>lower</u> after adjusting for non-completion                                 |
| Male:             |                              |       |                        |  |
| Associate degrees | >                            | >     | >                      | Returns even <u>greater</u> after adjusting for non-completion                               |
| Certificates      | <                            | >     | >                      | Lower <i>ex post</i> returns fully offset by higher completion and returns to non-completion |

Notes: Based on coefficients from Table 3.

results are not reported.) Of more interest is the comparison between those who intend to get a diploma and those who intend to get an Associate degree (regardless of their success). As shown in columns (3) and (4) of Table 3, the gains for Associate degrees are +11% [+5%] based on intent and +8% [+8%] based on goals. For females, at least, this result accords with those described above. When only looking at completers, the returns to an Associate degree over a diploma are +5%. After adjusting for probability of completion, the returns increase to 8–11%. Looking only at completers therefore under-states the returns to an Associate degree. However, for males, the opposite conclusion is evident: gains for Associate degree-holders are over-stated compared to gain for Associate degree enrollees.

As an alternative exposition, we estimated the predicted earnings for each group. We estimated the models in Table 2 to generate predicted earnings with a binary indicator for each award status controlling for covariates. These predicted earnings were quarterly earnings in 2011, i.e. seven to nine years after initial enrollment. Overall, these results showed *ex ante* returns to be considerably below *ex post* returns.<sup>10</sup>

Table 4 provides an overall summary of the results, comparing *ex ante* and *ex post* returns to diplomas to those of other awards. For females, those with associate degrees have higher returns than those with diplomas ( $R_C$ ) and those who intended to complete an associate degree also have higher returns ( $R_N$ ). The associate degree track also has a higher completion probability. The net result is that the returns to associate degrees are even greater after adjusting for non-completion. For certificates, the converse is true: relative to diplomas, the returns are lower for non-completers (as well as completers) and the completion rate is lower. Hence, the returns are even lower after adjusting for non-completion. For male students, the results for associate degrees are equivalent to those for females: returns are even greater when we adjust for non-completion. However, the results for certificates are distinctive. Males who complete a certificate earn less than males who complete a diploma. But males who do not complete a certificate earn more than those who do not complete a diploma. As well, the completion rate for certificates is higher than for diplomas. Given the sizes of the co-

efficients and probability, we calculate as per Eq. (2) above that the lower *ex post* returns are fully offset by the relatively higher returns to non-completion and the higher probability of completion. Overall, for males, we therefore find no difference in returns for those beginning a certificate versus those beginning a diploma.

#### 4.4. Labor market returns by subject

We now estimate returns by subject, separately identifying those for completers and the group who intended to complete. The subjects are collapsed from the CIP codes and follow the categories used by Dadgar, Venezia, Nodine, and Bracco (2013); the default subject is 'missing/other'. First, for each sample grouping we estimate the returns to subjects for students who have any award (degree, certificate or diploma) in that subject. All students who did not complete an award are grouped together. Therefore, all 'Award' coefficients are relative to students who got an award in a 'missing' subject. In the second specification, all students in each subject (labeled 'field') are grouped together regardless of whether they received an award or not, i.e. all non-completers are re-assigned to their field of study. Again, the default subject is 'missing/other' and so includes those who complete and do not complete. By comparing the two specifications, we are able to see the effect of including all students in a subject grouping versus only those students who pass their subject-specific award. We combine all students across award types; we do this partly to make the analysis manageable and partly because of sample size issues. However, to check the results we also restrict the sample just to students in the Associate degree programs.

Table 5 shows the returns across subjects using the two specifications ('Award' and 'Field'). For female students who receive awards, there are few earnings differences by subject (column 1 of Table 5). Most students who get an award get similar earnings, at least similar to the returns for students for whom the subject award is unknown. The clear exception is nursing, where earnings gains for nursing credentials are enormous (+207%); with an alternative comparison group, awards in health would probably also yield statistical significance. Unsurprisingly, as shown in the final row of column (1) of Table 5, non-completers have lower earnings than completers whose subject is unknown (by –20%).

<sup>10</sup> Details are not reported but are available from the authors.

**Table 5**

Returns to subjects: actual and predicted.

Average quarterly earnings gains in 2011 for 2002–2005 NCCCS cohorts.

|                           | Female              |                      |                                       |                     | Male                |                     |                                       |                     |
|---------------------------|---------------------|----------------------|---------------------------------------|---------------------|---------------------|---------------------|---------------------------------------|---------------------|
|                           | All Students        |                      | Associate degree students (predicted) |                     | All students        |                     | Associate degree students (predicted) |                     |
|                           | Award               | Field                | Award                                 | Field               | Award               | Field               | Award                                 | Field               |
| Humanities/Social Sci.    | −0.024<br>[0.131]   | 0.048<br>[0.050]     | 0.086<br>[0.180]                      | 0.087<br>[0.092]    | 0.304***<br>[0.101] | 0.143***<br>[0.055] | −0.085<br>[0.140]                     | −0.115<br>[0.078]   |
| Math/Science              | −0.485<br>[0.350]   | −0.104<br>[0.074]    | −0.410<br>[0.387]                     | −0.183<br>[0.171]   | 0.288<br>[0.285]    | 0.120<br>[0.097]    | −0.141<br>[0.303]                     | −0.145<br>[0.161]   |
| Inf. Science              | −0.049<br>[0.135]   | 0.005<br>[0.053]     | 0.090<br>[0.183]                      | 0.051<br>[0.094]    | 0.409***<br>[0.103] | 0.192***<br>[0.056] | 0.030<br>[0.142]                      | −0.087<br>[0.079]   |
| Engineering Mech.         | 0.039<br>[0.140]    | 0.111*<br>[0.056]    | 0.194<br>[0.186]                      | 0.202**<br>[0.096]  | 0.554***<br>[0.102] | 0.310***<br>[0.056] | 0.209<br>[0.140]                      | 0.082<br>[0.078]    |
| Health                    | 0.187<br>[0.130]    | 0.110**<br>[0.051]   | 0.394**<br>[0.180]                    | 0.262***<br>[0.092] | 0.644***<br>[0.105] | 0.265***<br>[0.060] | 0.298**<br>[0.143]                    | 0.091<br>[0.084]    |
| Nursing                   | 0.732***<br>[0.131] | 0.331***<br>[0.051]  | 1.022***<br>[0.180]                   | 0.717***<br>[0.092] | 0.979***<br>[0.106] | 0.331***<br>[0.059] | 0.692***<br>[0.144]                   | 0.267***<br>[0.085] |
| Mechanics                 | −0.238<br>[0.170]   | −0.152**<br>[0.075]  | −0.163<br>[0.277]                     | −0.075<br>[0.147]   | 0.390***<br>[0.101] | 0.179***<br>[0.055] | 0.140<br>[0.142]                      | 0.003<br>[0.080]    |
| Protective Services       | 0.151<br>[0.134]    | 0.018<br>[0.053]     | 0.219<br>[0.183]                      | 0.055<br>[0.094]    | 0.701***<br>[0.101] | 0.404***<br>[0.056] | 0.297**<br>[0.141]                    | 0.133*<br>[0.079]   |
| Construction              | −0.402<br>[0.273]   | −0.126<br>[0.093]    | −0.430<br>[0.794]                     | 0.045<br>[0.269]    | 0.266**<br>[0.105]  | 0.088<br>[0.058]    | 0.295**<br>[0.148]                    | 0.043<br>[0.087]    |
| Business/Marketing        | −0.015<br>[0.131]   | 0.029<br>[0.051]     | 0.144<br>[0.180]                      | 0.145<br>[0.092]    | 0.334***<br>[0.106] | 0.199***<br>[0.057] | −0.020<br>[0.144]                     | −0.039<br>[0.080]   |
| Education/Child Care      | −0.176<br>[0.131]   | −0.203***<br>[0.050] | 0.015<br>[0.180]                      | −0.042<br>[0.092]   | 0.280*<br>[0.145]   | 0.079<br>[0.059]    | −0.242<br>[0.205]                     | −0.254**<br>[0.101] |
| Transportation            | −0.273<br>[0.181]   | −0.245**<br>[0.102]  | 0.589**<br>[0.293]                    | 0.381*<br>[0.227]   | 0.232**<br>[0.109]  | −0.008<br>[0.064]   | −0.031<br>[0.319]                     | 0.030<br>[0.140]    |
| Cosmetology               | −0.201<br>[0.133]   | −0.199**<br>[0.052]  | 0.075<br>[0.184]                      | −0.002<br>[0.095]   | 0.278**<br>[0.123]  | 0.012<br>[0.064]    | −0.003<br>[0.163]                     | −0.157*<br>[0.089]  |
| Did not complete an award | −0.210<br>[0.130]   |                      | −0.019<br>[0.179]                     |                     | 0.326***<br>[0.099] |                     | −0.040<br>[0.139]                     |                     |
| Observations              | 96,221              | 96,221               | 48,885                                | 48,885              | 62,769              | 62,769              | 33,444                                | 33,444              |
| R-squared                 | 0.144               | 0.128                | 0.162                                 | 0.141               | 0.164               | 0.164               | 0.153                                 | 0.151               |

Notes: Log average quarterly earnings in 2011 as dependent variable. Cohorts of FTIC students from Fall 2002 to Summer 2005. Model includes: race (3), single parent, HS graduate, disability, enrollment age (2), experience, experience squared, financial aid EFC, grants, loans, and aid, college GPA first semester, college fixed effects. Robust standard errors in brackets.

\*\*\*  $p < 0.01$ ,

\*\*  $p < 0.05$ ,

\*  $p < 0.1$ .

Column (2) of Table 5 shows the pattern of returns for students who have been assigned to a field regardless of whether they completed an award or not. With this specification, the gains for nursing and health programs are substantially reduced (+37% and +12%, respectively). This diminution suggests that, although nursing programs are highly lucrative, they are less so when one accounts for the lower probability of completion. Moreover, new gaps emerge for the lowest-earning subjects: programs in Mechanics, Education/Child Care, Transportation and Cosmetology are now relatively less lucrative. Thus, given their relatively high failure rates, these subjects now appear to yield statistically significant penalties (compared to the unknown category). STEM subjects also show some low returns. The pattern is however not consistent, possible because many STEM students transfer to four-year colleges.

As shown in columns (3) and (4) of Table 5, the results are similar when the sample is restricted just to Associate programs. Again, nursing and health programs are the most lucrative but their premiums – albeit remaining positive –

are attenuated when we adjust for the probability of completing an Associate degree in these fields. Looking across the sample of students who are following a path toward an Associate degree, we find that adjusting for the likelihood of completion reduces subject-specific earnings gaps. For example, relative to those with associate degrees in unknown/missing subjects, returns to associate degrees in nursing/health are 278%/46% higher and the returns to all non-completers are 2% lower. When these non-completers are assigned to their predicted track, the returns to associate degrees in nursing/health over those in unknown subjects fall to 205%/30%. Although the gap remains very large, the *ex ante* returns to nursing [health] are one-third lower than the *ex post* returns for females.

For male students, the patterns in Table 5 are different because subject-specific enrollment patterns are very different by sex within community colleges. For the full sample of students there are high returns to Engineering Mechanics and Protective Services, as well as the health subjects. Although other subjects appear to be high-earning, they are not

**Table 6**

Returns to credits and program progression.

Average quarterly earnings gains in 2011 for 2002–2005 NCCCS cohorts.

|                                     | Female              |                     |                      | Male                |                    |                      |
|-------------------------------------|---------------------|---------------------|----------------------|---------------------|--------------------|----------------------|
|                                     | (1)                 | (2)                 | (3)                  | (1)                 | (2)                | (3)                  |
| <b>A. Full sample</b>               |                     |                     |                      |                     |                    |                      |
| Credits earned                      | 0.003***<br>[0.000] |                     | 0.004***<br>[0.000]  | 0.001***<br>[0.000] |                    | 0.002***<br>[0.001]  |
| Progression in program              |                     | 0.098***<br>[0.015] | –0.114***<br>[0.025] |                     | 0.040**<br>[0.019] | –0.089**<br>[0.033]  |
| Observations                        | 67,781              | 67,771              | 67,771               | 46,978              | 46,977             | 46,977               |
| R-squared                           | 0.131               | 0.130               | 0.131                | 0.169               | 0.169              | 0.169                |
| <b>B1. Only AA (pred.)</b>          |                     |                     |                      |                     |                    |                      |
| Credits earned                      | 0.000<br>[0.000]    |                     | –0.003*<br>[0.001]   | 0.000<br>[0.000]    |                    | 0.001<br>[0.001]     |
| Progression in program              |                     | 0.027<br>[0.025]    | 0.219**<br>[0.105]   |                     | 0.030<br>[0.027]   | –0.025<br>[0.097]    |
| Observations                        | 28,085              | 28,085              | 28,085               | 23,368              | 23,368             | 23,368               |
| R-squared                           | 0.136               | 0.136               | 0.136                | 0.155               | 0.155              | 0.155                |
| <b>B2. Only certificate (pred.)</b> |                     |                     |                      |                     |                    |                      |
| Credits earned                      | 0.004***<br>[0.001] |                     | 0.004***<br>[0.001]  | 0.005***<br>[0.001] |                    | 0.009***<br>[0.002]  |
| Progression in program              |                     | 0.128***<br>[0.021] | –0.001<br>[0.041]    |                     | 0.083**<br>[0.032] | –0.151***<br>[0.055] |
| Observations                        | 24,438              | 24,438              | 24,438               | 15,671              | 15,671             | 15,671               |
| R-squared                           | 0.138               | 0.137               | 0.138                | 0.199               | 0.198              | 0.200                |
| <b>B3. Only diploma (pred.)</b>     |                     |                     |                      |                     |                    |                      |
| Credits earned                      | 0.003***<br>[0.001] |                     | 0.004*<br>[0.002]    | –0.000<br>[0.001]   |                    | –0.006**<br>[0.003]  |
| Progression in program              |                     | 0.206***<br>[0.034] | –0.024<br>[0.139]    |                     | 0.031<br>[0.048]   | 0.343**<br>[0.155]   |
| Observations                        | 15,248              | 15,248              | 15,248               | 7,938               | 7,938              | 7,938                |
| R-squared                           | 0.140               | 0.140               | 0.140                | 0.174               | 0.174              | 0.175                |

Notes: Log average quarterly earnings in 2011 as dependent variable. Cohorts of FTIC students from Fall 2002 to Summer 2005. Sample restricted to students who did not complete an award. Model includes: race (3), single parent, HS graduate, disability, enrollment age (2), experience, experience squared, financial aid EFC, grants, loans, and aid, college GPA first semester, college fixed effects. Robust standard errors in brackets.

\*\*\*  $p < 0.01$ ,

\*\*  $p < 0.05$ .

when compared to the group of non-completers, who also have relatively high earnings compared to those who complete an award in the ‘missing field’ category. As for the female sample, the returns to all students within a given field are more closely clustered than are the returns only to completers within that field. Consistently, therefore, the greater earnings in some fields partly reflect the relative difficulty in program completion. For the sample of students in Associate degree programs, subject-specific variations in earnings are much lower. Few subjects show relatively high returns. Again, earnings gaps by subject award are much greater than by subject field. The greatest effect is for nursing: the very large relative earnings premiums for these awards are reduced by a factor of three when we include all the students who were predicted to be in a nursing program but who did not complete their award.<sup>11</sup>

<sup>11</sup> As for female students, the returns are reduced significantly for males. Relative to those with associate degrees in unknown/missing subjects, returns to associate degrees in nursing/health are 191%/29% higher and the returns to all non-completers are 6% [8%] lower. Re-assigning the non-completers reduces the returns to 127%/9%.

#### 4.5. Labor market returns for progressing in a program

Next we estimate the effects of students’ progression on their future earnings to see if students who have gotten further in their program do better in the labor market than those students who have accumulated general credits. As completers have progression fully, the sample is restricted to non-completing students. The estimates on progression are reported in Table 6. The full sample results are given in Panel A.

Model (1) estimates the returns to credits earned: for each credit, female [male] earnings are subsequently higher by 0.3 [0.1]%. These are significant gains, showing that dropouts still benefit from having more human capital.<sup>12</sup> Model (2) estimates the returns to progression towards an award: these returns are also positive. Students who progress further in their program have higher earnings.

However, because progression is strongly (although not perfectly) associated with more credits, it is not surprising

<sup>12</sup> In fact, these estimates are quite close to those of Jacobson et al. (2005), who find a gain of approximately 7% for males and 14% for females who attend community colleges in Washington state (assuming the average number of credits completed at 27).



that it is positively associated with earnings. Of interest here is whether, conditional on their credit accumulation, a student who has progressed further towards completion will have higher earnings. Model (3) tests for the net effect of progression: it shows that, among non-completers, controlling for total credits earned, students who have progressed further in their program of study classes earn less than students who have taken general credits. The penalty for progression is similar for both female and male non-completers: students who progress twice as far as the average non-completer does earn 4–5% less (and the effect of credits has not attenuated in this broader model).

Panel B of Table 6 shows the results split by predicted award. It may be that the penalty from progressing but not completing varies across award: vocational or licensing certificates may be especially sensitive, with perhaps no reward to an uncompleted licence in mechanical engineering for example. However, the results in Panel B show mixed patterns, with negative, positive and null returns to progression.

To our knowledge there is no prior evidence to corroborate this finding that progression is not superior to credit accumulation. It may be that there are diminishing returns to subject matter (over general courses) or that taking upper-level subject-specific courses is inconvenient. It may be that these individuals have had less time in the labor market. Or, as suggested by a reviewer, it may be that strictly vocational courses (e.g. phlebotomy in nursing) have little value outside that vocation and such a vocation is closed off to all non-completers. Overall, this finding does not support theories that students who progress further are more motivated or higher ability (conditional on them not completing their program). Instead, it suggests a ‘high-stakes’ scenario where, unless a student is going to graduate, their labor market outcomes would be improved if they dropped out of a program as quickly as possible.

## 5. Discussion and conclusion

Most community college students do not complete a credential, even among those who enter college indicating this is their intention. There is substantial current interest in boosting college completion rates. Along with the increased emphasis on completion, there has also been increased interest in programs of study, and on getting students to adopt and stick to a program of study relatively early in their college career, with the view that this will boost completion (Dadgar et al., 2013; Jenkins & Cho, 2012). Yet college faculty, administrators, and staff do not typically have a good sense of what credentials their students are pursuing, and what amount of progress those students have made toward that goal (Gardenhire-Crooks, Collado, & Ray, 2006; Grubb, 2006). Furthermore, there is little information on the economic value of student progress – as opposed to the wealth of information on the economic value of award completion. There is very little information in the literature on their expected return to pursuing a program, adjusted for the probability of completing it, which is what these *ex ante* results provide.

This information is needed for two reasons. First, if colleges are to raise completion rates, colleges will need a good sense of where their students are in their programs in order

to help them move along their program pathway. Even in the case in which a student has explicitly informed the college as to what credential she is pursuing, her actual coursework may not correspond well to her declared major. Second, given low completion rates, the economic output of community colleges is driven by the value of human capital accumulated by these non-completers, not just that of the relatively few completers.

The challenge is to classify students who do not complete an award. Both stated preference and revealed preference methods may be appropriate. Although given students’ lack of understanding and the complexity of programs offered by colleges, we argue for a greater focus on revealed preference methods to interpret student progress. Certainly, different classification systems – by algorithm, intent or goal – yield very different enrollment patterns.

More importantly, these classifications make a substantial difference to earnings patterns. Not only does completion of a particular award yield a specific pay-off, but failure to complete does as well. For example, we find that across the non-completers failing a certificate program appears to carry the heaviest penalty. Combining the analysis for completers and failures also makes a difference. Completion of an Associate degree yields higher earnings than completion of a diploma. But after we account for the probability of completion, this premium actually increases.

By contrast, the results by subject show that returns are compressed once we adjust for the probability of completion. For example, the returns to nursing are extremely high but the completion rate is very low such that the *ex ante* returns – although still substantial – are significantly diminished. In general, therefore, the widespread finding that subject of study matters a lot to returns to college – although correct – is probably an overstatement of the true differential (Webber, 2014). From a policy perspective, it is important for students to take into account not only the returns to a program, but the probability of actually completing that program.

Finally, we find that progression *per se* does not lead to higher earnings for students who do not complete (even as it demonstrably does for students who complete their program). This result, which is surprising but not implausible, may help explain why so many students do drop out. After taking the relatively valuable general courses at college, these students may have insufficient motivation, time, or funds for (or find inconvenient to attend) the upper-level courses which are no more valuable in the labor market but that are required for graduation. Viewed from a social perspective, this finding suggests that college personnel in charge of programs might face a high-stakes dilemma: forcing students into a structured program is beneficial if they complete; if they don’t complete, it is more efficient for them to accumulate general credits. Regardless of how rationally students make these decisions, however, considerably more attention should be paid to the human capital obtained by non-completers.

## Appendix

Fig. A1., Table A1.

### Assignment of Students to Programs

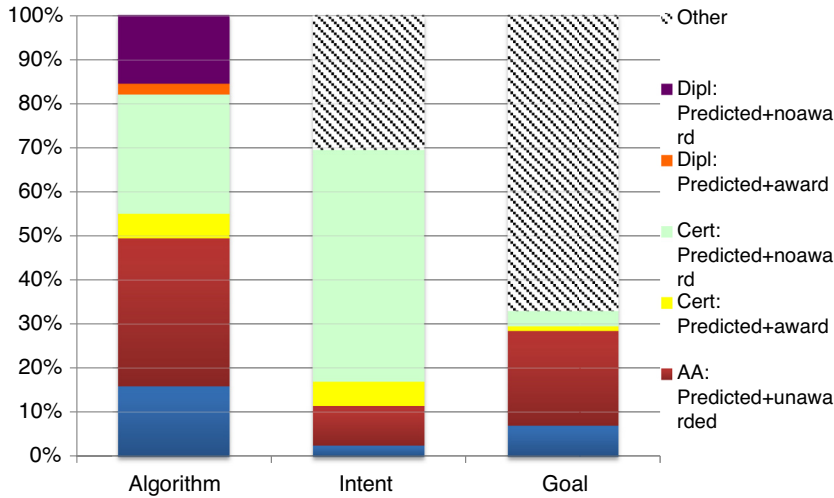


Fig. A1. Assignment of students to programs.

Table A1

Descriptive statistics for sample with valid earnings data.

|   | Female      | Male        |
|---|-------------|-------------|
| <b>Awards</b>                             |             |             |
| Associate degree                          | 0.22        | 0.16        |
| Certificate                               | 0.04        | 0.07        |
| Diploma                                   | 0.03        | 0.02        |
| <b>Students predicted per award</b>       |             |             |
| Associate degree                          | 0.51        | 0.53        |
| Certificate                               | 0.29        | 0.31        |
| Diploma                                   | 0.20        | 0.16        |
| <b>Intentions</b>                         |             |             |
| Associate degree (intended and award)     | 0.05        | 0.01        |
| Associate degree (intended but no award)  | 0.14        | 0.05        |
| Certificate (intended and award)          | 0.04        | 0.06        |
| Certificate (intended but no award)       | 0.61        | 0.50        |
| Intended other                            | 0.29        | 0.39        |
| <b>Goals</b>                              |             |             |
| Associate degree (goal and award)         | 0.11        | 0.07        |
| Associate degree (goal but no award)      | 0.26        | 0.21        |
| Certificate (goal and award)              | 0.00        | 0.00        |
| Certificate (goal but no award)           | 0.04        | 0.04        |
| Goal other                                | 0.58        | 0.67        |
| Program match                             | 0.30        | 0.33        |
| <b>Credits earned</b>                     |             |             |
| Associate degree – award                  | 76.2 (23.1) | 71.8 (22.1) |
| Certificate – award                       | 36.4 (23.1) | 26.9 (20.2) |
| Diploma – award                           | 61.5 (20.1) | 52.5 (20.1) |
| Associate degree – predicted but no award | 27.2 (23.7) | 24.7 (23.1) |
| Certificate – predicted but no award      | 10.6 (13.5) | 8.7 (11.6)  |
| Diploma – predicted but no award          | 19.6 (19.7) | 15.8 (17.5) |
| <b>Progression</b>                        |             |             |
| Associate degree – predicted but no award | 0.35 (0.30) | 0.32 (0.30) |
| Certificate – predicted but no award      | 0.32 (0.39) | 0.28 (0.37) |
| Diploma – predicted but no award          | 0.31 (0.31) | 0.29 (0.32) |
| Ln (Quarterly Earnings) <sup>a</sup>      | 8.16 (1.20) | 8.35 (1.24) |
| Total observations                        | 96,221      | 62,769      |

<sup>a</sup> Non-zero earnings observations only.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.econedurev.2015.09.004](https://doi.org/10.1016/j.econedurev.2015.09.004).

### References

Agan, A. (2013). *Disaggregating the returns to college*. Princeton University Working paper.

Altonji, J. G., Blom, E., & Meghir, C. (2012). Human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, 4(1), 185–223.

Andrews, R. J., Li, J., & Lovenheim, M. F. (2012). *Quantile treatment effects of college quality on earnings: Evidence from administrative data in Texas*. Cambridge, MA: National Bureau of Economic Research (NBER working paper no. 18068). Retrieved from [www.nber.org/papers/w18068](http://www.nber.org/papers/w18068).

Bahr, P. R. (2010). The bird's eye view of community colleges: A behavioral typology of first-time students based on cluster analytic classification. *Research in Higher Education*, 51, 724–749.

Bailey, T., Alfonso, M., Scott, M., & Leinbach, T. (2004). *The educational outcomes of occupational postsecondary students*. Washington, DC: U.S. Department of Education, National Assessment of Vocational Education.

Bailey, T., & Belfield, C. R. (2012). Community college occupational degrees: are they worth it? In L. Perna (Ed.), *Preparing today's students for tomorrow's jobs in metropolitan America*. Philadelphia, PA: University of Pennsylvania Press.

Bailey, T., Jenkins, D., & Leinbach, T. (2006). *Is Student Success labeled institutional failure? Student goals and graduation rates in the accountability debate at community colleges*. Working paper no. 1. New York: Community College Research Center, Teachers College, Columbia University.

Belfield, C. R., & Bailey, T. (2011). The benefits of attending community college: a review of the evidence. *Community College Review*, 39(1), 46–68.

Bound, J., Lovenheim, M. F., & Turner, S. (2010). Why have college completion rates declined? An analysis of changing student preparation and collegiate resources. *American Economic Journal: Applied Economics*, 2(3), 129–157.

Brand, J. E., & Xie, Y. (2010). Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *American Sociological Review*, 75(2), 273–302.

Carneiro, P., Heckman, J. J., & Vytlačil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, 101(6), 2754–2781.

Caruana, R., & Niculesu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on machine learning* (pp. 161–168).

- Choy, S. P., & Horn, L. J. (1992). *A guide to using postsecondary transcript data and an overview of course taking in less-than-four-year postsecondary institutions*. Berkeley, CA: National Center for Research in Vocational Education.
- Compton, J., Laanan, F. S., & Starobin, S. (2010). Career and technical education as pathways: Factors influencing postcollege earnings of selected career clusters. *Journal of Education for Students Placed at Risk*, 15, 93–113.
- Dadgar, M., & Weiss, M. J. (2012). Labor market returns to sub-baccalaureate credentials: How much does a community college degree or certificate pay? (CCRC working paper no. 45). New York, NY: Columbia University, Teachers College, Community College Research Center.
- Dadgar, M., Venezia, A., Nodine, T., & Bracco, K. R. (2013). *Providing structured pathways to guide students toward completion*. San Francisco: WestEd.
- Desjardins, S. L., Ahlburg, D. A., & McCall, B. P. (2002). A temporal investigation of factors related to timely degree completion. *Journal of Higher Education*, 73, 555–581.
- Doyle, W. R. (2009). The effect of community college enrollment on BA completion. *Economics of Education Review*, 28, 199–206.
- Hand, D. J., & Yu, J. (2001). Idiot's Bayes—Not so stupid after all? *International Statistical Review*, 69, 385–398.
- Hossler, D., Shapiro, D., Dundar, A., Ziskin, M., Chen, J., Zerquera, D., et al. (2012). *Transfer and mobility: A national view of pre-degree student movement in postsecondary institutions*. Herndon, VA: National Student Clearinghouse Research Center Retrieved from <http://www.nscresearchcenter.org/signaturereport2>.
- Gardenhire-Crooks, A., Collado, H., & Ray, B. (2006). *A whole another world: Students navigating community college*. New York, NY: MDRC.
- Grubb, W. N. (2006). Like, what do I do now? The dilemmas of guidance counseling. In T. Bailey, & V. Morest (Eds.), *Defending the community college equity agenda* (pp. 195–222). Baltimore, MD: Johns Hopkins University Press.
- Jacobson, L., & Mohker, C. (2009). *Pathways to boosting the earnings of low-income students by increasing their educational attainment*. Washington, DC: Hudson Institute Center for Employment Policy.
- Jacobson, L., LaLonde, R., & Sullivan, D. G. (2005). Estimating the returns to community college schooling for displaced workers. *Journal of Econometrics*, 125(1–2), 271–304.
- Jenkins, D., & Cho, S. (2012). *Get with the program: Accelerating community college students' entry into and completion of programs of study*. Working paper no. 32. New York: Community College Research Center, Teachers College, Columbia University.
- Jepsen, C., Troske, K., & Coomes, P. (2014). The labor market returns to community college degrees, diplomas and certificates. *Journal of Labor Economics*, 32, 95–121.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the international joint conference on artificial intelligence* (pp. 1137–1143).
- Liu, Y. T., Belfield, C. R., & Trimble, M. (2014). Estimating the medium term returns to labor market credentials: Evidence for North Carolina. *Economics of Education Review*, 44, 42–55.
- Liu, V. Y. T., Belfield, C. R., & Trimble, M. J. (2015). The medium-term labor market returns to community college awards: Evidence from North Carolina. *Economics of Education Review*, 44, 42–55.
- Melguizo, T., Kienzl, G., & Alfonso, M. (2011). Comparing the educational attainment of community college transfer students and four-year college rising juniors using propensity score matching methods. *Journal of Higher Education*, 82, 265–291.
- Mitchell, T. (1997). *Machine learning*. New York: McGraw Hill.
- Radford, A. W., Berkner, L., Wheelless, S. C., & Shepherd, B. (2010). *Persistence and attainment of 2003–04 beginning postsecondary students: After 6 years*. Washington, DC: U.S. Department of Education, National Center for Education Statistics (NCES 2011–151).
- Rish, I. (2001). An empirical study of the naive bayes classifier. In *Proceedings of the IJCAI-01 workshop on empirical methods in artificial intelligence* (pp. 41–46).
- Romano, R. M., Losinger, R., & Millard, T. (2011). Measuring the cost of a college degree: A case study of a SUNY community college. *Community College Review*, 39(3), 211–234.
- Stuart, G. R. (2009). A benefit/cost analysis of three student enrollment behaviors at a community college: Dropout, transfer and completion of an associate's degree/certificate (Doctoral dissertation). Retrieved from <http://etd.ohiolink.edu>.
- Webber, D. A. (2014). The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors. *Labour Economics*, 28, 14–23.
- Wu, X., Kumar, V., Quinlan, R. J., Ghosh, J., Yang, Q., Motoda, H., et al. (2008). Top 10 algorithms in data mining. *Knowledge Information Systems*, 14, 1–37.
- Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics*, 29, 301–348.