Proceedings of the 2008 College of Liberal Arts Forum on Faculty-Student Research

March 26, 2008

Rochester Institute of Technology

A Path Dependent Model of Course Selection

Arthur Johnston

Economics

Abstract

Classes that college students take in their first few semesters can dramatically influence their later choice of classes. A student who finds she is interested in Science and bored by Liberal Arts in her first semester is more likely to take Science classes in the future. This trend is self-reinforcing (or has positive feedback): each following semester the student's probability of taking a Science class increases and the probability of taking a Liberal Arts class decreases. This paper models such a path dependent situation using a non-Markovian stochastic process. Using this model also allows us to examine different ways colleges can affect the student's education, both through courses required to graduate and requiring a balanced course load each semester.

Introduction

At every point in life people make choices based on the expected utility of those choices. Since people rarely have perfect information about the future, these choices are based on previous experiences. A specific example would be high school students entering college who are in a difficult position because their choice of college major and classes they will take all have a major influence on their future. Unfortunately high school students by definition have relatively little experience compared to later in life. What little experience students do have may influence them into choosing a major or classes that, if they had more experience they would not choose. This can lead to a less than optimal situation for both the student and society. Although much research has been done exploring why students choose a major, much less has been done to develop a model of how students choose individual classes.[1,2]

Many factors influence what classes college students take in a given semester. They include courses needed to graduate, other courses taken that semester, and fields of interest. Courses needed to graduate include both specific courses required and broader requirements such as "four Liberal Arts classes." Students also balance their course load each quarter so that they are not taking all difficult classes in any one quarter. A student's areas of interest are subjects that the student is curious about.

Modeling the third influence, a student's subject of interest, is the subject of this paper. The model chosen is path dependant: decisions are not independent but are to some degree influenced by previous decisions.

As students take classes in a subject, the probability of them taking more of that same subject in the future increases. The first part of this paper presents the model and some results that can be obtained with it. Using the same model, the second part of the paper also presents a way of influencing each student's class choice.

The Basic Model of Path-Dependent Course Selection

If no other conditions were imposed on them, most students will take classes only in the subjects that interest them. In the simplest model, if a scientifically minded student has a choice between taking one course in either a Liberal Art or a Science, they will take Science. If the model does not allow for changes of interest this model will lead to a student only ever taking classes in one field.

However students do not take classes in only one field. Occasionally the Science student will take a Liberal Art and vice versa. This can simply be modeled by having the Science student's chance of taking a Science class be a probability, P_S, such that $0< P_S<1$. Obviously this means that the probability of taking a Liberal Art is now $P_L=1- P_S$ if there are only Liberal Arts and Science classes.

However, this model does not allow for a student's interest to change over time; it also does not allow a student who, at the start of college, is interested in both Science and Liberal Arts to specialize once he finds his niche. To model a student's interest in a subject, at any given point in time, we look at both initial interest and previous classes taken. This can be modeled:

$$P_S = \frac{I_S + C_S}{(I_S + C_S + I_L + C_L)} \qquad P_L = \frac{I_L + C_L}{(I_S + C_S + I_L + C_L)}$$

Where I is the initial interest in the subject and C is the number of classes taken in the field. The special case where I_L and I_S are equal to 1 is a simple case of Polya's urn and the student is equally likely to take any distribution of classes taken.[3]

In most cases, students are not equally interested in Liberal Arts and Science, $I_S \neq I_L$ and have taken no courses $C_S = C_L = 0$. Since students are initially more interested in one subject over the other they are more likely to specialize in the subject of their initial interest. A simple example would be if a student is initially twice as interested in Science as she is Liberal Arts, $I_L = 1$ and $I_S = 2$, and has taken no classes, $C_S = C_L = 0$. If she had taken one class per semester for two semesters, after the first semester there are two possibilities.

1. The student chose a Science class in the first semester which has 2/3 probability.

2. The student chose a Liberal Art class in the first semester which has 1/3 probability

After the first semester either C_S or C_L is now one so P_S and P_L have now changed to either

$$P_S = \frac{2+1}{(2+1+1+0)} \qquad P_L = \frac{1+0}{(2+1+1+0)}$$

if a Science class was chosen and

$$P_S = \frac{2+0}{(2+0+1+1)} \qquad P_L = \frac{1+1}{(2+0+1+1)}$$

if a Liberal Arts class was taken. This means for the second semester there are four possibilities:

1. The student chose a Science class in the first semester which has 2/3 probability. The student then chooses a Science class their second semester which has probability 3/4. This combination has a probability of 1/2.

2. The student chose a Science class in the first semester which has 2/3 probability. The student then chooses a Liberal Arts class their second semester which has probability 1/4. This combination has a probability of 1/6.

3. The student chose a Liberal Art class in the first semester which has 1/3 probability. The student then chooses a Science

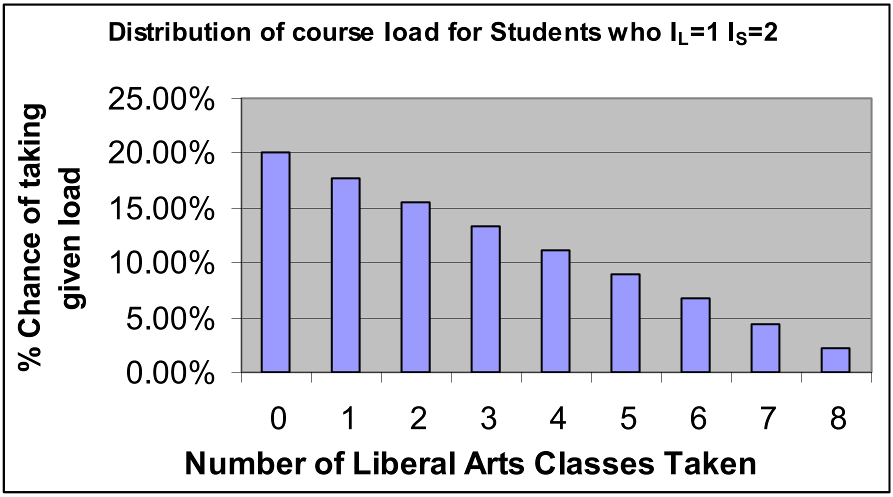
class their second semester which has probability 1/2. This combination has a probability of 1/6.

4. The student chose a Liberal Art class in the first semester which has 1/3 probability. The student then chooses a Liberal Arts class their second semester which has probability 1/2. This combination has a probability of 1/6.

Possibilities two and three are actually the same- the student takes one Science class and one Liberal Arts class; only the order is different. Combining possibilities two and three the probability distribution for this student is

50% chance of taking two Science classes
33% chance of taking one Science class and one Liberal Arts class
16% chance of taking two Liberal Arts classes

A better example is if a student is initially twice as interested in Science as he is in Liberal Arts, $I_L = 1$ and $I_S = 2$. If he takes one class per semester for eight semesters he is likely to overspecialize in Science. In the chart below we can see that the probability of a student taking no Liberal Arts classes is twenty percent.



The results are more revealing if we break them into three groups. The first group consists of specializing in Science, that is zero, one and two Liberal Arts. The second is a balanced course load, three, four and five Liberal Arts. The final group is specializing in Liberal Arts, six, seven and eight Liberal Arts. Fifty-three percent of the distribution falls into the Science specialization group while only one third of the students are in the balanced course load group.

Obviously students normally take more than one class per semester. The probability distribution changes and initial differences in interest are more influential. For example, if as in the previous example a student is initially twice as interested in Science as she is in Liberal Arts, $I_L = 1$ and $I_S = 2$, and she takes two classes per semester after one semester there are four scenarios:

1. The student chose a Science class as the first class. The student then chooses a Science class as the second class. This combination has a probability of (2/3)*(2/3)=4/9

2. The student chose a Science class as the first class. The student then chooses a Liberal Art class as the second class. This combination has a probability of (2/3)*(1/3)=2/9

3. The student chose a Liberal Art class as the first class. The student then chooses a Science class as the second class. This combination has a probability of (1/3)*(2/3)=2/9

4. The student chose a Liberal Art class as the first class. The student then chooses a Liberal Art class as the second class. This combination has a probability of (1/3)*(1/3)=1/9

Possibilities two and three are actually the same, so after one semester there are three possible situations. Likewise, with the second semester there are eight possibilities but some of them are actually the same. The final probability distribution for this student is:

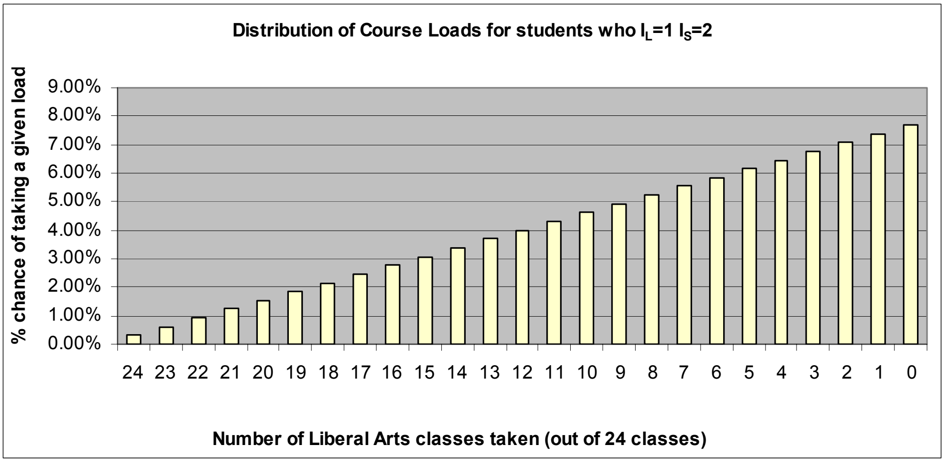
4% chance of taking four Liberal Arts classes
12.44% chance of taking three Liberal Arts classes and one Science class

24.88% chance of taking two Liberal Arts classes and two Science classes

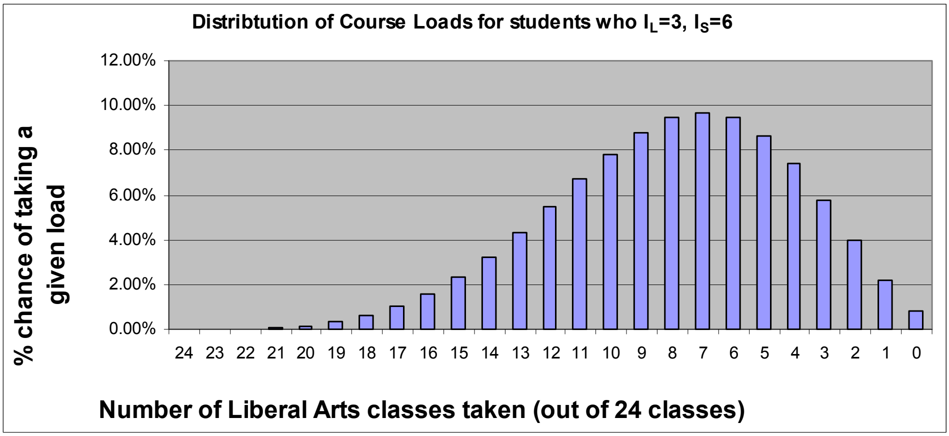
30.22% chance of taking one Liberal Arts class and three Science classes

28.44% chance of taking four Science classes

A more realistic model would enable the student to take three classes a semester over the course of eight semesters, or four years. If this student were again initially twice as interested in Science as Liberal Arts, $I_L = 1$ and $I_S = 2$, the probability distribution would be:



This distribution is essentially the same as taking eight classes over eight semesters. A student who was more strongly focused but was still twice as interested in Science classes than Liberal Arts might be represented with, $I_L = 3$ and $I_S = 6$.

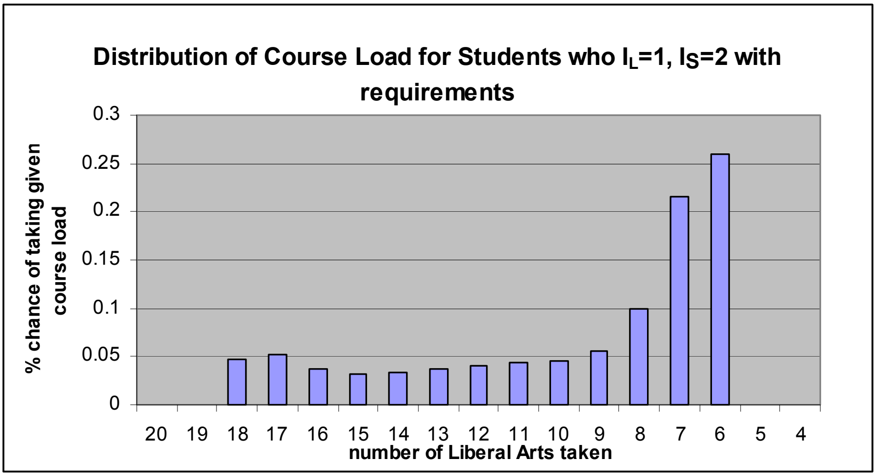


The extremes for this more focused student are much smaller than those of the less focused student. This is because the initial interests are great enough that no single quarter, including the first, can completely change his focus. Therefore, the stronger a student's initial interest in a subject the less likely he or she is to be swayed by one or two classes.

Influencing Course Selection

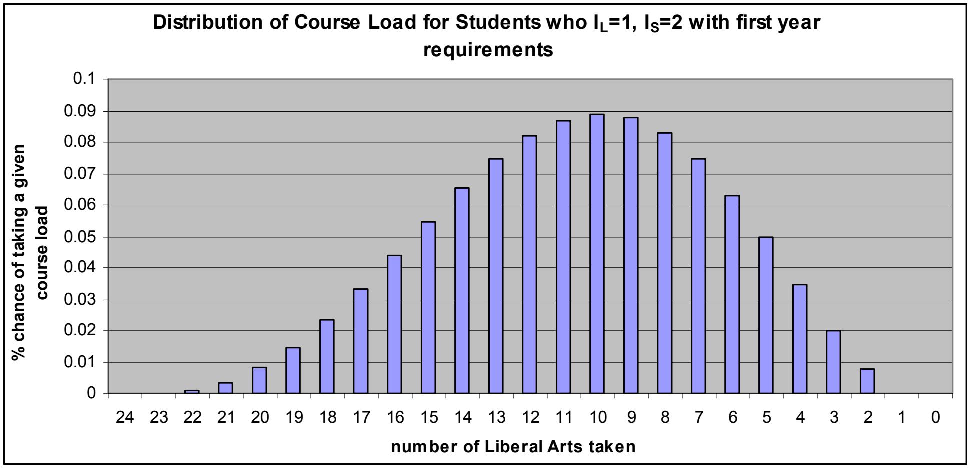
Such overspecialization into a single subject area obviously leads to less studying in other areas. It is easy to imagine a situation where a student only takes classes in his subject area of interest, to the point where he completely avoids other classes. Obviously some college administrators view this negatively and try to make well rounded students. A fairly common way to do this is for students to have distribution requirements for graduation. This means that they must take a certain number of classes outside of their major.

The requirements placed in this model were that students showed little foresight. A student would take the class load she would normally take until she is close to graduation. She then takes all the classes she is required to graduate. This leads to a curved distribution since the 'tails' have been cut off. Looking at someone twice as interested in Science as Liberal Arts, $I_L = 1$ and $I_S = 2$, taking three classes a semester for eight semesters, but adding the requirement that she takes at least six Liberal Arts and six Sciences, final C_S=>6 and final C_L=>6, produces the following graph:



Since some course loads are disallowed (a student cannot take fewer than six Liberal Arts, nor can they take more than eighteen, if they only take twenty-four classes total), if a student were likely to end up on either tail of the graph he is instead funneled towards the taking the minimum of a subject required to graduate. Although this produces a better rounded student it also has the effect of dictating at least a quarter of a student's classes. This requires students to take classes they are less interested in and curtails their choices.

A better way of ensuring more well rounded students is to make the student want to take a more diverse course load. This can be done by becoming less restrictive in graduation requirements and more selective in freshman requirements. Instead of requiring a student take six Liberal Arts and six Science classes before graduating, instead require him or her to take two Liberal Arts and two Sciences freshman year. That is at T=2 C_S=>2 and final C_L=>2. This yields:



This distribution of probabilities leads to many more well rounded students. There are still some students who only take classes in one field of interest, but they are an insignificant part of the population. The distribution is still skewed in favor of taking more Science classes than

Liberal Arts but much less so than any of the other distributions, including the distribution that requires six of both Liberal Arts and Sciences.

Conclusion

Each student's choice of classes, especially in the first two years in college, has a dramatic effect on later class choice and later life decisions. Since these first few classes and the students' initial interest have the most effect, a path dependant model is appropriate. Such a model shows how the first few classes a student takes can dramatically influence later life decisions. If college administrators wish to influence students into taking a more balanced combination of courses there are two major ways to do so: imposing graduation requirements, which students will complete grudgingly or having freshman course distribution requirements that are less intrusive and actually affect the students' interests.

References

1. Berger, Mark C. 1988. "Predicted Future Earnings and Choice of College Major**."**
Industrial and Labor Relations Review 41(3): 418-429

2. Botelho, A., Pinto, L.C. 2004. "Students' Expectations of the Economic Returns to College Education: Results of a Controlled Experiment." *Economics of Education Review* 23(6): 645-653

3. W. Brian Arthur, Yuri M. Eermoliev, and Yuri M. Kaniovski. 1987. "Path-Dependent Processes and the Emergence of Macrostructure." *European Journal of Operational Research* 30: 294-303.