When They Will Pay: Understand Deposit Decisions in College Admission

## Introduction

Motivation Recruitment of new students is a challenging task of enrollment management for many higher education institutions, since they need to meet revenue goals, ensure academic ability and promote diversity of the student body (Adams-Johnson et al. 2019; Maldonado, Armelini, and Guevara 2017). The recruitment process includes several general stages, searching and answering inquries of propsective applicants, reviewing application materials and making admission decisions, encouraging admitted students to pay deposits, and assisting with students’ matriculation (Litten et al. 1983). Although matriculation is the final outcome, the deposit stage is really the last focus of the Admission Office, because most students paying deposit will eventually matriculate. Other offices/departments in an institution pay attention to deposit situation too. The Budget Office would like to use deposits paid to estimate the tuition revenue from new students each fall. The Budget Office need to update budgetary planning in the spring, if the deposits paid are far away from the admission targets. Similarly, the Department of Housing needs to work on the potential shortage of dorm space in the spring, if too many deposits are paid. As a result, the Admission Office track deposits usually by week in the spring and desire to understand admitted students’ decision to pay deposits, not only whether they will pay, but also when they will pay, so they can adjust marketing and recruitment stratigies promptly (DesJardins 2002; Goenner and Pauls 2006).

Summary of previous studies Previous studies have provided valuable insights to understand students’ enrollment decisions and to assist with marketing and recruitment stratigies. DesJardins (DesJardins 2002) used predictive models to segment admitted students, and he suggested enrollment managers to focus on student groups who have enrollment probabilities close to 0.5. Maldonado et al. (Maldonado, Armelini, and Guevara 2017) developed nested logit models to predict admitted students’ enrollment probabilities, and the decision makers utilized the models to allocate resources for admission activities. Goenner and Pauls (Goenner and Pauls 2006) predicted the enrollment probabilities of inquirers to help University of North Dakota allocate recruitment efforts by geographic area. Braunstein et al. (Braunstein, McGrath, and Pescatrice 1999) focused on the effects of student financial factors on enrollment decisions of admitted students, so the particular institution could understand how various kinds of financial aid affected students differently from different socio-economic backgrounds. Johnson (Johnson 2019) studied various factors associated with the enrollment decisions of out-of-state students. He identified multiple potential destinations of the students and provided insights on why the students accepted their offer of admission or chose another instittion.

Difference from previous studies and contribution of this study Similar to many previous studies, we investigate factors to affect admitted students’ decisions, but we focus on when they will pay deposits rather than whether they will matriculate. So instead of the cross-sectional methods used in previous studies, we use event history analysis to model the deposit decisions, which is a popular tool to model students’ journey from matriculation to graduation (Chen and Hossler 2017; Gross, Torres, and Zerquera 2013; Zhan, Xiang, and Elliott III 2018). There are several challenges

Second, factors may have time-varying effects on deposit decisions. For example, students may be encouraged to pay deposits after attending a campus event, but the effect would decay over time. We develop Piecewise exponential models to address this challenge. Different from the proportional hazard models used in the previous studies on graduation, Piecewise exponential models can accomodate not only factors with time-varing values but also time-varing effects. This technique is especially helpful to evaluate the deadline effect, i.e., students who would like to attend an instuition have to pay the deposits by the deadline, so the driving force to pay in the last period is much higher than that in the initial period. Piecewise exponential models however are more prone to overfitting, comparing to proportional hazard models. We thus introduce Bayesian Hierarchical framework to the models. The framework provides balance between overfitting and underfitting, so true time-varing effects are not exaggerated by noise in observed data or ignored by model assumptions. Third, it is hard to evaluate students’ interests to admission offers received. The more interests students have where they are admitted to, the sooner they may pay the deposits. Hence it is helpful to understand their deposit decisions if we know whether they desire to attend an institution. We address this challenge with including student activities as surrogates of students’ interests to the admission offers, in addition to the factors usually included in previous research such as demographic characteristics, socio-economic background, high school academic performance and student financials. More specifically, we investigate the effects of factors related to students’ activities on the admission websites (Lee and Fu 2008; Neumann, Tucker, and Whitfield 2019) and campus visits, such as time spent on major finder pages, time spent on student finance pages, whether delaying to review the admission decision, and whether attending any admission events.

Summary of results In this paper, we provide insights to understand students’ deposit decisions. As expected, the most important factor is the deadline which is May 1st each year at the University of Delaware. Students are most likely to pay deposits in the last three days before deadline, more likely to pay in April, and less likely beforehand. Besides the deadline effect, our results indicate that many factors have time-varying effects on deposit decisions, such as gender, institutional grant offered, Pell eligibility and early campus events. The other factors’ effects do not change over time, such as loan borrowed, racial ethnicity and time spent on admission websites. These findings can help the Admission Office to understand when admitted students’ will pay deposits and thus help with recruitment efforts.

## Theoretical Backgound and Pratical Application

### Admission Funnel

From institution perspective, an admission funnel include students from six stages, prospects, inquiriers, applicants, admittants, depositors and matriculants (Litten et al. 1983). In the prospect stage, the Admission Office searches for high school students who might be interested to the institution. In the inquiry stage, the Admission Office communicates with students who expressed interest, attempts to further increase their interest, and encourages them to apply. In the applicant stage, the Admission Office notifies students with incomplete application forms, process and review completed applications. In the admit stage, the Admission Office makes decisions to offer admission, puts students on a wait list, or rejects applications. In the deposit stage, the Admission Office interacts with admitted students in campus tours and other programs, and Student Financial Services offer financial aid packages, in order to encourge the students to accept the offers. In the enrollment stage, the Admission Office collaborates with other offices to help with new student orientation, course registration and on-campus residency. There are several important rates to monitor in the admission funnel, conversion rate, selection rate, yield rate and melt rate (**ruffalonl?**). Conversion rate is the proportion of applicants from inquiries. Selection rate is the proportion of admits from applicants. Yield rate is the proportion of deposits from admits. Melt rate is the proportion of matriculants from deposits. With melt rate typically closed to 1 or 100% at the University of Delaware, the deposit stage or yield rate largely determines the number of new students we will have.

### College Choice

From student perspective, students’ college choice has three broad stages, college aspiration formation, search and application, and selection and attendance (Hossler and Gallagher 1987; Paulsen 1990). Students decide whether to go to college in the stage of college aspiration formation. The stage typically starts in early childhood and finishes in high school, but can last beyond high school. Various factors can influence students’ decisions in the stage, such as family background (Stage and Hossler 1989; Carpenter and Fleishman 1987), encouragement of teachers (Portes and Wilson 1976) and counselors (Conklin and Dailey 1981), and students’ academic aptitude and achivement (Tuttle 1981; Davies and Kandel 1981). Students decide which colleges to apply in the search and application stage. Most students start to create a list of institutions to apply in junior years of high school and finish the applications in the senior years (Gilmour Jr et al. 1981). In addition to information from parents, counselors and peers, institutions reach out to the students with college publication such as guidebooks and invite them to campus events (Goenner and Pauls 2006). After being admitted, students decide which college to attend in the final stage. Students make decisions based on their demographic background, socio-economic status, academic preparation, and institution characteristics such as cost, financial aid, academic programs, academic reputation and location (DesJardins 2002; Goenner and Pauls 2006).

Many studies have investigated students’ enrollment decisions in the selection and attendance stage with their individual characteristics and preferences. DesJardins (DesJardins 2002) used a logistic regression model to predict the enrollment probabilities of students who were admitted to a public institution in the Midwest in Fall 1999 and Fall 2001. The variables include students’ demographic and socio-economic background, high school characteristics, when they applied, and personal intension. The students were grouped into deciles according to the predicted enrollment probabilities. He suggested that it was more effcient to target the “fence-sitting” groups than the groups with very high enrollment probabilities. Goenner and Pauls (Goenner and Pauls 2006) used logistic regression models with Bayesian model average techniques to predicted the enrollment probabilities of 15,827 inquirers who were interested to attend the University of North Dakota in Fall 2003, to help allocate recruitment efforts by geographic areas. They investigated the effects of inquiry contact methods, geographic factors, geodemographic factors, academic factors and some interaction terms. They suggested to concentrate recruitments effort to geographic areas where enrollment probabilities are high. Johnson (Johnson 2019) focused on 42,950 out-of-state students who were admitted to a public research univiersity from Fall 2012 to Fall 2016 and used mixed multinomial models to investigate their enrollment decisions. He used National Student Clearinghouse data to identify whether the students chose to attend one of five destinations: the study institution, another out-of-state public institution, an in-state public institution, a private institution, or a 2-year college. The factors include demographic characteristics, high school information, family background, financial aid offered and admitted academic discipline. He found that students were more likely to attend the institution if family incomes were 85,000 or higher, a family member was an alumnus or an alumna, graduated from a feeder high school, being offered higher merit scholarship, or borrowed more loans. A surpursing finding was Pell-eligible students’ enrollment decisions were not affected by grants offered in financial aid packages. Maldonado et al. (Maldonado, Armelini, and Guevara 2017) used nested logit models to predict enrollment decisions of 25,325 prospective students to four Bachelor’s programs in a small private Chilean university. The three possible outcomes of the students were applied, admitted but not enrolled, and admitted and enrolled. They argued that the hierarchical models are necessary, because the last two outcomes were more similar than the first outcome and should be grouped together. Three groups of factors were included, marketing efforts from the institution, students’ socioeconomic background, and stated preferences of the students. They found that on-campus activities and talks at secondary schools were more effective than career fairs to encourage enrollment, male students were more likely to attend engineering and law programs, and students’ online activities and stated preferences indicated their interest to attend the institution or individual programs. Braunstein et al. (Braunstein, McGrath, and Pescatrice 1999) used logistic regression to model the enrollment decisions of 7,104 students admitted to Iona College in three academic years. They included three groups of variables, demographic and social background, academic achievment and academic preparation, and financial aid. They found that the demographic and social background did not affect the students’ enrollment decisions, but the financial aid had positive effect. The enrollment probability increased between 1.1% and 2.5%, for every additional (dollar)1,000 offered. Within the financial aid offered, loans borrowed had more influence than grants offered, and work study had the least influence.

### Conceptual Framework

Our conceptual framework is developed from the theory of college choice (Chapman 1979; Hossler, Braxton, and Coopersmith 1989) and previous studies on students’ enrollment decisions. Students decide whether to accept admission offers, according to economic, sociological, and psychological factors (Paulsen 1990). Based on the theory of human capital (Becker 2009), the economic factors help students evaluate whether an institution provides the highest potential benefit versus cost. From sociological perspective, students choose the best institution for status attainment (Hossler, Schmit, and Vesper 1999), and the choice is affect by their sociological background (Johnson 2019). The psychological factors impact students’ opinions on institutional environment and climate, and students decide whether it is a good fit to the institution (Paulsen 1990). We hypothesis that these factors not only affect whether students would like to pay deposits, but also when they would like to pay. For admitted students, they are willing to pay deposits sooner than later if they are admitted to an institution with the best potential investment, status attainment, and/or student-institution fit.

With the guidance of the college choice theory and previous studies, we form the following variable pool for the three kinds of factors. The economic factors include variables related to financial aid, institutional financial aid offered, loan borrowed, Pell eligibility, and time spent on student finance service (SFS) websites. Financial aid affects students’ choice by reducing the cost of attendence (Braunstein, McGrath, and Pescatrice 1999), so we hypothesize that students are more likely to pay deposits with more financial aid. Students could evaluate which institution provides the best financial aid packages till the last minute, so we hypothesize that the effect of financial aid will be more obvious when deposit deadline is approaching. The sociological factors include students’ demographic characteristics, socio-economic status and high school academic performance, i.e., gender, ethnicity, expected family contribution (EFC), and HS GPA. They reflect the influence on students’ choice from parents, peers, counselors and teachers in high school (Johnson 2019). We hypothesize that they have various effects on students’ deposit decisions and they may affect the decisions in difference paces.

The psychological factors reflect students’ desire to attend the institution (Paulsen 1990), which include the period to receive admission offer, whether admitted to the Honor College, whether applied major and admitted major are different, whether attended recruitment events, whether postpone to review admission decision, and time spent on major finder websites (Maldonado, Armelini, and Guevara 2017). We hypothesize that students are more likely to pay deposits in early periods, if they have high interests to the institution. We hypothesize that it is an impulsive force to encourage students to pay deposits when students know they are admitted, so we include the period to receive admission offer. On the other hand, students who postpone or ignore the admission decision may have low desire to attend. Being admitted into the Honor College is affirmation to students’ academic achievement and could raise students’ interest to pay deposits. Students and the Admission Office usually communicate with each other, before being admitted to a major which is different from the applied major. Being willing to be admitted to a different major is a good indictor that they desire to attend the institution, no matter which academic program to attend. Other students may pay more attension to the academic programs, because being able to attend the desired program is important to determine whether it is a good student-institution fit. We thus include time spent on major finder websites to reflect this kind of students’ consideration. Lastly, attending the recruitment events shows students’ interest to learn more about the institution and thus are more likely to accept admission offers.

## Data and Variables

The study institution, University of Delaware (UD), is a public research university (Carnegie classification: R1) with a population of about 18,000 undergraduate students. The Admission Office of provided the admission data of 58,426 applicants who intended to matriculate as first-time first-year students in Fall 2020, Fall 2021 and Fall 2022. We track the deposit decisions of admitted out-of-state students from February 1 to the deposit day or May 1 (deposit deadline). We use February 1 as the starting point, because the Admission Office has made most admission decisions by then. We divide the time between February 1 and May 1 into eight periods, February, March 1 to March 15, March 16 to March 31, April 1 to April 7, April 8 to April 14, April 15 to April 21, April 22 to April 28, and April 29 to May 1. Table 1 shows the number of observations and deposits by period each year. For each year, the number of observations increases from period 1 to period 3 or period 4, because some students were admitted after February. The number of observations decreases afterwards, because we stop to track students who paid deposits. The numbers of deposits increase in April, especially after April 21, indicating the deadline effect on students’ deposit decisions.

Table 1. Numbers of Observations and Deposits by Period Each Year | |2020 | |2021 | |2022 | | | ——| ——|——-|——-|——-|——-|——-| |Period |N |Deposit|N |Deposit|N |Deposit| |1 - February |15866 |308 |17785 |226 |18019 |263 | |2 - March 1 to March 15 |16457 |247 |18330 |173 |18736 |241 | |3 - March 16 to March 31 |16673 |378 |19257 |381 |19304 |385 | |4 - April 1 to April 7 |16385 |303 |19110 |357 |19465 |316 | |5 - April 8 to April 14 |16112 |384 |18891 |434 |19289 |454 | |6 - April 15 to April 21 |15779 |416 |18459 |550 |18839 |461 | |7 - April 22 to April 28 |15739 |620 |17983 |757 |18412 |962 | |8 - April 29 to May 1 |15129 |481 |17226 |576 |17467 |685 |

Table 2 describes the definitions of the variables in the models. The depedent variable is whether a student paid deposit by May 1. The independent variables include the the economic factors, sociological factors and psychological factors listed in the concepture framework. The three variables related to student financials, institutional grant, Loan and EFC, are standardized by dividing the monetary amounts over COA. Two of them have time-varying values. Table 3 shows their descriptive statistics by period. The average offered instutional grant rate increases from 0.190 in period 1 to 0.205 in period 8. Given the COA for out-of-state students is about $$53K, the average increase is about $$800. Similarly, The average loan borrowed increases from 0.071 in period 1 to 0.082 in period 8. The average increase is about $$580. The rest variables have time-independent values. Table 4 shows their descriptive statistics. For binary categorical variables, table 4 shows the numbers and percentages of students in the category associated with value 1. For example, we admitted 7077 Pell eligible students in the three years, which is 12.1% of all admitted students. For numeric variables, table 4 shows their mean, standard deviation, maximum and minimum values. For example, students spent 290.31 minutes on the SFS websites on average, the standard devivation is 740.08. The most time spent on the SFS websites is 29531 minutes, and some students do not browse any SFS website.

[Insert Table 2 Here]

[Insert Table 3 Here]

[Insert Table 4 Here]

$ $

## Statistical Model

We develope Bayesian hierarchical piecewise exponential models to understand students’ deposit decisions. Piecewise exponential models are a kind of discrete event history analysis with constant harzard function within each discrete time interval (DesJardins et al. 1994; Friedman 1982). In this study, an event occurs if a student pays deposit. Otherwise, the admitted student is censored or “survives” from the desire to pay deposit. We assume that the desire accumulates over time, and the rate changes in different time periods according to the student’s attributes. Equation (1) defines the logarithm of the harzard (rate) function to be the sum of two parts, baseline hazard and a linear combination of student attributes, where denotes the value of variable j for student i in a period and \_{j}[Period] denotes the effect of the variable j in the period. The cumulative hazard is the product of the harzard function and the period length , and the logrithm of survival function is defined to be the negative cumulative hazard as shown in Equation (2). And finally, Equation (3) defines the probablity of the student i paying deposit in a period () to be 1 - the survival function , i.e., the probability of the student not survive in the period. If we think the hazard function reflects the driving force to encourage the student to pay deposit, cumulative hazard is the accumlated force over time, i.e., impulse in term of physics. If a variable has positive effect on the deposit decision, then the higher the variable value, the higher driving force, the higher impulse, and thus the higher probablity to pay deposit according to three equations.

(1)

(2)

(3)

$

We introduce the Bayesian hierarchical framework to estimate the unknown coefficients and in Equation (1). In a Bayesian analysis, our initial uncertainty, known as a prior distribution, is modelled using the language of probability distributions (McElreath 2020). For the baseline hazard , we model the initial uncertainty using normal distributions as shown in Equation (4). The is the maximum likehood estimates for . We choose to be 0.1 to make it a strong prior distribution. That is to say, when the prior distributions are updated to posterior distributions using data observed, the data must show strong evidence to let the posterior distribution deviate from prior distribution. The hierarchical structure is setup for estimating . At the higher level, the variables have time-independent effects , i.e., the average effects over all periods. We assume are normally distributed with means being 0 as shown in Equation (5). This is to reduce overfitting for the higher level effects, because the observed data need to show enough support for non-zero parameter estimates. At the lower level, the variables have time-varying effects . We assume are normally distributed with means being as shown in Equation (6). This is to reduce overfitting for the lower level effects, because the observed data need to show enough support for to deviate from . The uncertainty of our assumptions is controlled by the standard deviations, and , respectively. We model the uncertainty of and with expoential prior distributions with rate parameter 0.5 as shown in Equation (7) and (8). In summary, if the observed data provide enough support, some posterior distributions of the higher level effects will deviate from the zero-mean prior distributions, and some posterior distributions of the lower level effects will deviate from the prior distributions centered at the higher level effects, so we really have some variables with time-varying effects on the deposit decisions.

(4)

(5)

(6)

(7)

(8)

## Results and Discussion

The posterior distributions of the effects are estimated using the DynamicHMC package [] in Julia (version 1.7.2). The effects of variables are estimated using three Markov chains. Each chain contains 1,000 samples, after a series of warmup steps to find proper sizesize for the “No-U-Turn Sampler” (Betancourt 2017; Hoffman, Gelman, et al. 2014).

The results are interpreted in terms of log-log deposit probabilities. According to formulas (1) to (3),

(9)

### Fall 2020

All variables have important effects on yield decisions, and they all have time-varying effects.

[Insert Table 5 Here]

In a Bayesian analysis, the time-varying effects are easily visualized.

[Insert Figure 1 Here]

Time-varying effects

Time-independent effects

Not important variables

Comparison among the three years Common patterns Covid effect Change of admission policy

## Reference

Adams-Johnson, Susan, Jeff Cranmore, Anna MJ Holloway, and Joel D Wiley. 2019. “Higher Education Recruitment in the United States: A Chronology of Significant Literature.” *Journal of Educational Administration and History* 51 (3): 213–38.

Becker, Gary S. 2009. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago press.

Betancourt, Michael. 2017. “A Conceptual Introduction to Hamiltonian Monte Carlo.” *arXiv Preprint arXiv:1701.02434*.

Braunstein, Andrew, Michael McGrath, and Donn Pescatrice. 1999. “Measuring the Impact of Income and Financial Aid Offers on College Enrollment Decisions.” *Research in Higher Education* 40 (3): 247–59.

Carpenter, Peter G, and John A Fleishman. 1987. “Linking Intentions and Behavior: Australian Students’ College Plans and College Attendance.” *American Educational Research Journal* 24 (1): 79–105.

Chapman, Randall G. 1979. “Pricing Policy and the College Choice Process.” *Research in Higher Education* 10 (1): 37–57.

Chen, Jin, and Don Hossler. 2017. “The Effects of Financial Aid on College Success of Two-Year Beginning Nontraditional Students.” *Research in Higher Education* 58 (1): 40–76.

Conklin, Mary E, and Ann Ricks Dailey. 1981. “Does Consistency of Parental Educational Encouragement Matter for Secondary School Students?” *Sociology of Education*, 254–62.

Davies, Mark, and Denise B Kandel. 1981. “Parental and Peer Influences on Adolescents’ Educational Plans: Some Further Evidence.” *American Journal of Sociology* 87 (2): 363–87.

DesJardins, Stephen L et al. 1994. “Studying the Determinants of Student Stopout: Identifying" True" from Spurious Time-Varying Effects.”

DesJardins, Stephen L. 2002. “An Analytic Strategy to Assist Institutional Recruitment and Marketing Efforts.” *Research in Higher Education* 43 (5): 531–53.

Friedman, Michael. 1982. “Piecewise Exponential Models for Survival Data with Covariates.” *The Annals of Statistics* 10 (1): 101–13.

Gilmour Jr, Joseph E et al. 1981. “How High School Students Select a College.”

Goenner, Cullen F, and Kenton Pauls. 2006. “A Predictive Model of Inquiry to Enrollment.” *Research in Higher Education* 47 (8): 935–56.

Gross, Jacob PK, Vasti Torres, and Desiree Zerquera. 2013. “Financial Aid and Attainment Among Students in a State with Changing Demographics.” *Research in Higher Education* 54 (4): 383–406.

Hoffman, Matthew D, Andrew Gelman, et al. 2014. “The No-u-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *J. Mach. Learn. Res.* 15 (1): 1593–623.

Hossler, Don, John Braxton, and Georgia Coopersmith. 1989. “Understanding Student College Choice.” *Higher Education: Handbook of Theory and Research* 5: 231–88.

Hossler, Don, and Karen Gallagher. 1987. “Studying Student College Choice: A Three-Phase Model and the Implication...-SuperSearch Powered by Summon.” *College and University* 62: 201–21.

Hossler, Don, Jack Schmit, and Nick Vesper. 1999. *Going to College: How Social, Economic, and Educational Factors Influence the Decisions Students Make*. JHU Press.

Johnson, Iryna Y. 2019. “Destinations of Admitted Out-of-State Students: A Case of One Institution.” *Research in Higher Education* 60 (3): 315–37.

Lee, Chu-Hui, and Yu-Hsiang Fu. 2008. “Web Usage Mining Based on Clustering of Browsing Features.” In *2008 Eighth International Conference on Intelligent Systems Design and Applications*, 1:281–86. IEEE.

Litten, Larry H et al. 1983. “Applying Market Research in College Admissions.”

Maldonado, Sebastián, Guillermo Armelini, and C Angelo Guevara. 2017. “Assessing University Enrollment and Admission Efforts via Hierarchical Classification and Feature Selection.” *Intelligent Data Analysis* 21 (4): 945–62.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Chapman; Hall/CRC.

Neumann, Nico, Catherine E Tucker, and Timothy Whitfield. 2019. “Frontiers: How Effective Is Third-Party Consumer Profiling? Evidence from Field Studies.” *Marketing Science* 38 (6): 918–26.

Paulsen, Michael B. 1990. *College Choice: Understanding Student Enrollment Behavior. ASHE-ERIC Higher Education Report No. 6.* ERIC.

Portes, Alejandro, and Kenneth L Wilson. 1976. “Black-White Differences in Educational Attainment.” *American Sociological Review*, 414–31.

Stage, Frances K, and Don Hossler. 1989. “Differences in Family Influences on College Attendance Plans for Male and Female Ninth Graders.” *Research in Higher Education* 30 (3): 301–15.

Tuttle, Ron. 1981. “A Path Analytic Model of the College Going Decision.”

Zhan, Min, Xiaoling Xiang, and William Elliott III. 2018. “How Much Is Too Much: Educational Loans and College Graduation.” *Educational Policy* 32 (7): 993–1017.