# Modeling Student Scheduling Preferences in a Computer-Based Testing Facility

**Matthew West**
Mechanical Science and Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
mwest@illinois.edu

**Craig Zilles**
Computer Science
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
zilles@illinois.edu


*L@S'16*, April 25–26, 2016, Edinburgh, Scotland.

DOI string from ACM form confirmation

## Abstract
When undergraduate students are allowed to choose a time slot in which to take an exam from a large number of options (e.g., 40), the students exhibit strong preferences among the times. We found that students can be effectively modelled using *constrained discrete choice theory* to quantify these preferences from their observed behavior. The resulting models are suitable for load balancing when scheduling multiple concurrent exams and for capacity planning given a set schedule.

## Author Keywords
student modeling; asynchronous exams; discrete choice theory; capacity planning; computerized testing

## ACM Classification Keywords
I.6.5 [Simulation and Modeling]: Model Development; H.1.2 [Information Systems]: User/Machine Systems

## Introduction
The Computer-Based Testing Facility (CBTF) [4] at the University of Illinois at Urbana-Champaign (UIUC) is a testing facility that allows students to self-schedule time slots for mid-term and final exams using the PrairieLearn [3] online problem-solving system.

In Fall 2015 the CBTF conducted about 2000 exams per

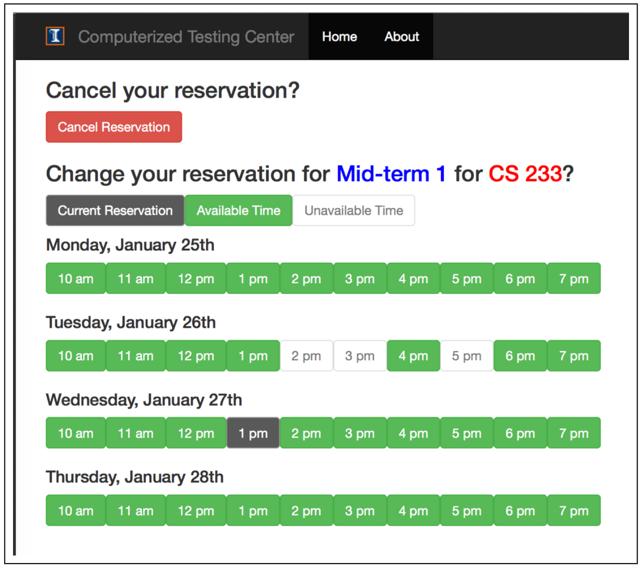


**Figure 1:** The student exam sign-up interface for the CBTF (Computer-Based Testing Facility).

week with 9 different large-enrollment courses scheduling an average of 6 total distinct exams per week. Each exam has a block of 3 to 5 days during which students are free to schedule their individual exam time. Exams start on the hour, typically 10am to 10pm daily, giving about 30 to 60 possible time slots that a student can choose from for each exam. However, the testing facility can only accommodate 45 students simultaneously, so there are significant constraints on student choice. Scheduling is handled via a self-service online system (the *CBTF Scheduler*) where students can sign up for any time slot that currently has seats available, and they are also able to delete or reschedule their future reservations. Figure 1 shows the student scheduling interface.

For planning purposes it is desirable to have a predictive model of student preferences and scheduling behavior. For example, different exams need to be spaced out over the course of a week to balance load across the days, and usage needs to be forecast so that more proctors can be scheduled during busy periods. Data from Fall 2015 shows strong and consistent patterns in student scheduling preferences and behavior, indicating that predictive models should exist.

## Constrained discrete choice models

To model student choices of time slots, we use the framework of *discrete choice theory* [2] from economics, which models choices that people make from a set of discrete alternatives. In particular, we use discrete choice models *with capacity constraints* [1], because each time slot in the CBTF can only accommodate a maximum of 45 students simultaneously.

We use $n$ to index *decisions*, which are students choosing time slots, $i$ to index time *slots*, $h$ to index the *hour-in-week* of a time slot, and $r$ to be the number of *days remaining* within the period of a given exam. Using these variables, we define four binary arrays:

$$\begin{aligned} a_{ni} &\in \{0,1\} \text{ is whether slot } i \text{ was available for decision } n \\ y_{ni} &\in \{0,1\} \text{ is whether slot } i \text{ was chosen for decision } n \\ x_{ih} &\in \{0,1\} \text{ is whether slot } i \text{ is at hour-in-week } h \\ w_{nir} &\in \{0,1\} \text{ is whether slot } i \text{ for decision } n \text{ has} \\ &\qquad r \text{ days remaining in the exam period.} \end{aligned}$$

The desirability of time slot $i$ for the student making decision $n$ is modeled by its *utility* $V_{ni}$. We will model this by assuming the utility of a time slot depends on which hour in the week it falls (that is, the hour of day and the day in week) and also how many days are remaining in the exam period. We expect that students will prefer certain hours and days (e.g., they might prefer late afternoon on weekdays) and that they will prefer to do exams close to the end of each exam period. Our model is described by parameters $\beta_h$ and $\lambda_r$ as follows:

$$\begin{aligned} V_{ni} &= \text{utility of slot } i \text{ for decision } n \\ &= \sum_{h=1}^{N_\mathrm{h}} \beta_h x_{ih} + \sum_{r=1}^{N_\mathrm{r}} \lambda_r w_{nir} \\ \beta_h &= \text{utility of hour-in-week } h \\ \lambda_r &= \text{utility of } r \text{ remaining days in the exam period.} \end{aligned}$$

Constrained discrete choice theory uses random utility models [2] to derive the probability $p_{ni}$ that decision $n$ chooses slot $i$ to be:

$$p_{ni} = \frac{a_{ni} e^{V_{ni}}}{\sum_{j=1}^{N_\mathrm{i}} a_{nj} e^{V_{nj}}}.$$

Given a set of actual choices in the array $y_{ni}$, we can determine the best-fit model parameters $\beta_h^*$ and $\lambda_r^*$ using maximum-likelihood estimation. This gives:

$$\begin{aligned} L &= \text{log-likelihood of the set of actual choices} \\ &= \sum_{n=1}^{N_\mathrm{n}} \sum_{i=1}^{N_\mathrm{i}} y_{ni} \log p_{ni} \\ \beta^*, \lambda^* &= \operatorname*{argmax}_{\beta,\lambda} L(\beta, \lambda). \end{aligned}$$

## Results

Using data from the CBTF at UIUC in Fall 2015 we found the maximum likelihood parameters shown in Figures 2 and 3. These parameters show that students prefer Thursdays and Fridays (but not Friday evenings), they prefer

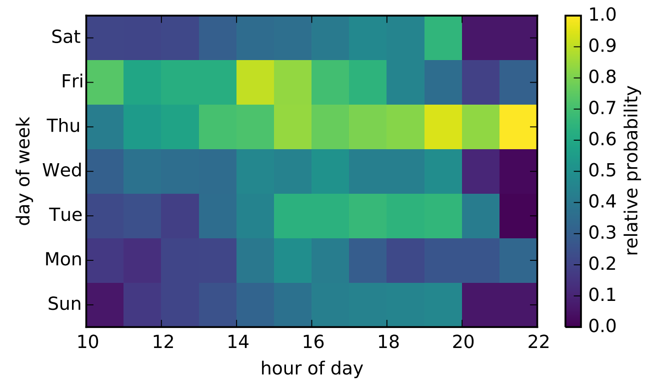


**Figure 2:** The relative probability (proportional to $\exp \beta_h^*$) of a student choosing a time slot as a function of the hour-in-week $h$.

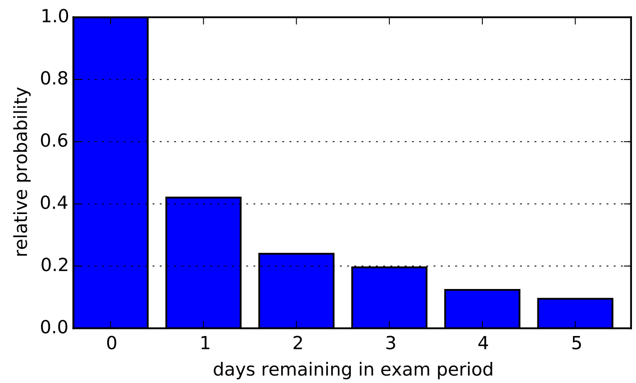


**Figure 3:** The relative probability (proportional to $\exp \lambda_r^*$) of a student choosing a time slot as a function of days remaining $r$.

| | Model | Uniform |
|---|---|---|
| Median absolute error | 3.7% | 9.6% |
| Mean absolute error | 7.2% | 14.5% |
| Root mean square error | 11.8% | 21.3% |

**Table 1:** Prediction errors from the discrete choice model ("Model") and a naive uniform prediction ("Uniform"). Errors are given as percentages of maximum capacity (45 students). The model-based predictions have about half the error of the uniform.

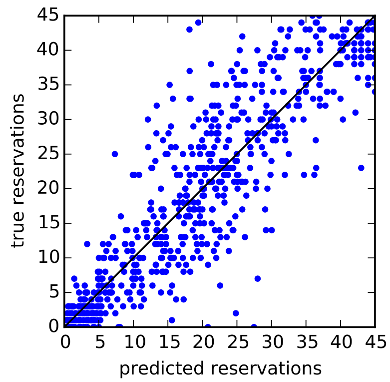


**Figure 4:** The predicted number of reservations per time slot using the discrete choice model, versus the true numbers. The 45° black line represents a perfect predictor.

afternoon time slots, and they very much prefer to do the exam on the last possible day.

The predicted versus true number of students per time slot are shown in Figure 4, which have errors as shown in Table 1. The discrete choice model generally performs well, with well under 10% mean absolute error in the number of students who choose any given time slot.

## Conclusions and future work

We fitted a constrained discrete choice model for student scheduling preferences in a Computer-Based Testing Facility (CBTF) and found that the model predicted time slot usages with a mean absolute error of 7.2%, indicating that it is a practical model for planning purposes. Future work includes full cross-validation studies and extensions to per-course parameters and per-student modeling.

## Acknowledgements

MW acknowledges support from NSF DUE-1347722.

## REFERENCES


1. A. de Palma, N. Picard, and P. Waddell. 2007. Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region. *Journal of Urban Economics* 62, 2 (2007), 204–230. DOI: http://dx.doi.org/10.1016/j.jue.2007.02.007

2. K. E. Train. 2009. *Discrete Choice Methods with Simulation* (second ed.). Cambridge University Press.

3. M. West, C. Zilles, and G. Herman. 2015. PrairieLearn: Mastery-based Online Problem Solving with Adaptive Scoring and Recommendations Driven by Machine Learning. In *Proceedings of the American Society for Engineering Education (ASEE) 2015 Annual Conference*. DOI: http://dx.doi.org/10.18260/p.24575

4. C. Zilles, R. T. Deloatch, J. Bailey, B. B. Khattar, W. Fagen, C. Heeren, D Mussulman, and M. West. 2015. Computerized Testing: A Vision and Initial Experiences. In *Proceedings of the American Society for Engineering Education (ASEE) 2015 Annual Conference*. DOI: http://dx.doi.org/10.18260/p.23726