Identification of At-Risk Students and Analyses of Stopout & Re-enrollment at an Hispanic-Serving Institution to Develop Interventions for Improving Degree Attainment: A Survival Model Approach

The University of Texas at El Paso
Denise Carrejo
Bereket Weldeslassie
Roy Mathew

Center for Institutional Evaluation, Research and Planning

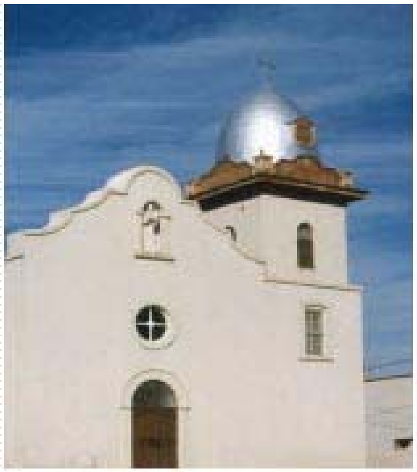
CIERP gratefully acknowledges the support of
Lumina Foundation for Education

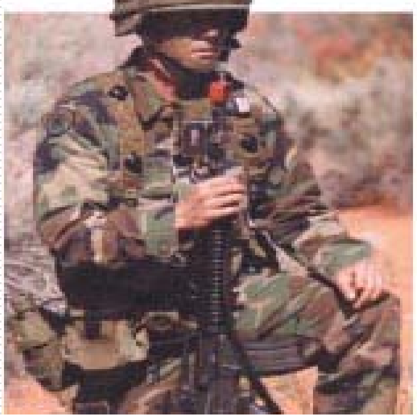
Presentation Outline

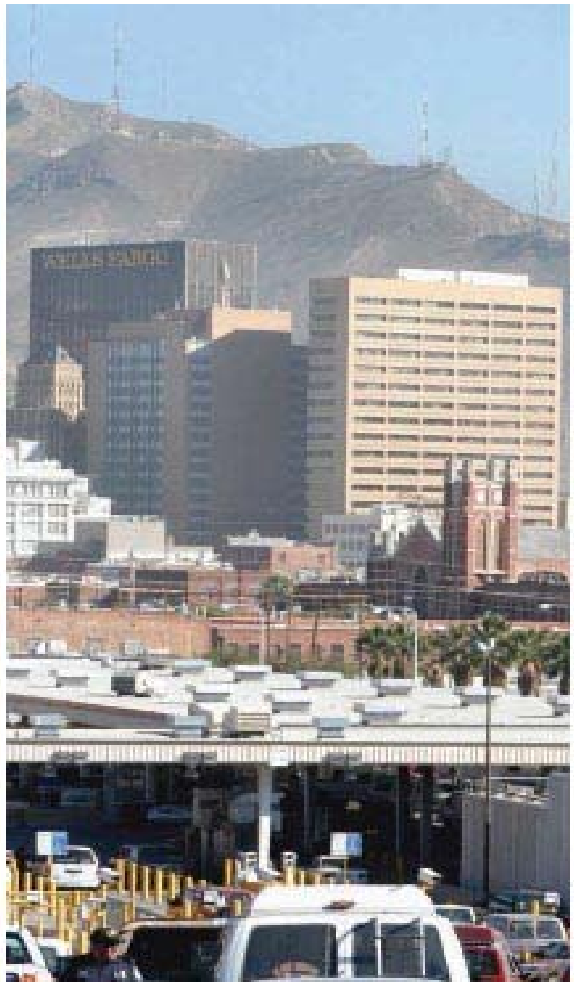
- UTEP's Context
- Factors that affect departure and reenrollment
- Factors that affect graduation within six years
- Risk groups and associated trends
- Summary, implications and next steps

UTEP's Context

El Paso, Texas







- El Paso County-3rd poorest large county in the US[1]
- Population: 734,000
- 81% Hispanic
- Border community, very dynamic flow of residents and students across the border
- Limited educational opportunities

[1] 2005 American Community Survey of the US Census

UTEP Demographics

Enrollment by Race/Ethnicity

	N	%
White Non-Hispanic	2,124	10.5%
Black Non-Hispanic	547	2.7%
Hispanic	**14,826**	**73.6%**
Asian/Pacific Islander	226	1.1%
Am. Indian or Alaskan	44	0.2%
International*	2291	11.4%

*includes Mexican Nat'l. students

Total Enrollment by Residence

	N	%
El Paso County	**16705**	**83.9%**
New Mexico	243	1.2%
Mexico	1801	8.9%
Other Int'l	431	2.1%

Percent of financial aid awardees with family income of *$20,000 or less*: 43%

Percent of UTEP students with reported family income of *$20,000 or less*: 33%

<u>Nationally</u>:

% of students with family income of less than $20,000 at large public research (doctoral) universities: 10%[1]

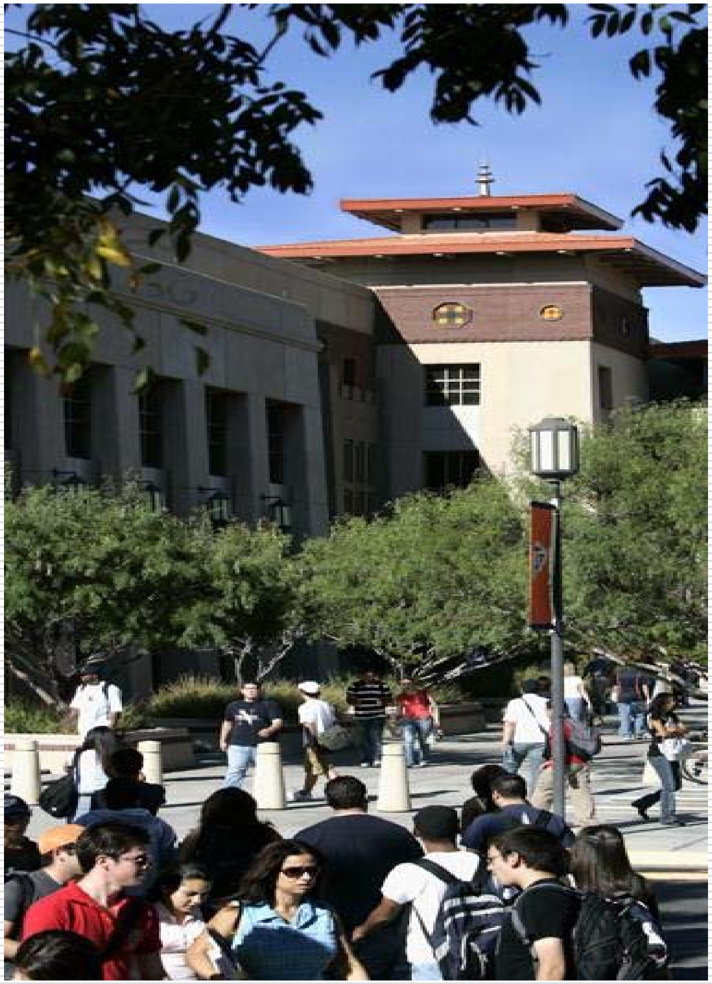
% of students with family income of less than $20,000 at small &mid-sized private colleges and universities: 12%[1]

% of students with family income less than $20,000 at community colleges: 29%[2]

[1]Council of Independent Colleges: http://www.cic.edu/makingthecase/data/access/income/index.asp

[2]Lumina Foundation Focus, Fall 2005, p.5

UTEP's Efforts over last 20 years



- President Diana Natalicio has rededicated UTEP's mission to ensure the widest possible access to all students from the region, and to focus on serving the El Paso area
- Institution has made efforts to ensure success at all levels of the pipeline, with impressive results
  - K-12
  - Admissibility/ Affordability
  - Student engagement

UTEP's Institutional Successes

- College of Engineering—identified as the top engineering school for Hispanics by *Hispanic Business Magazine*. UTEP "is changing the face of engineering and producing highly trained graduates heavily recruited by the industry's leading companies".[1]
- NSSE and the American Association for Higher Education identified UTEP as one of 20 colleges and universities that was "unusually effective in promoting student success". [2]
- One of only six NSF Model Institutions for Excellence in the nation for its success in creating educational opportunities for non-traditional students.
- Rankings:
  - 3rd in the US for granting baccalaureate degrees to Hispanics
  - 6th in the US for granting Master's degree to Hispanics

1989 UTEP Alumnus Danny Olivas, NASA astronaut flew on the shuttle Atlantis in June 2007

[1]NSSE Institute for Effective Educational Practice, Project DEEP Final Report, p. 4

[2]*Hispanic Business*, September 2007

Despite its Success, UTEP has More Work to Do

Six year, FTFT graduation rate:

- Fall 2001 cohort → 29.4%

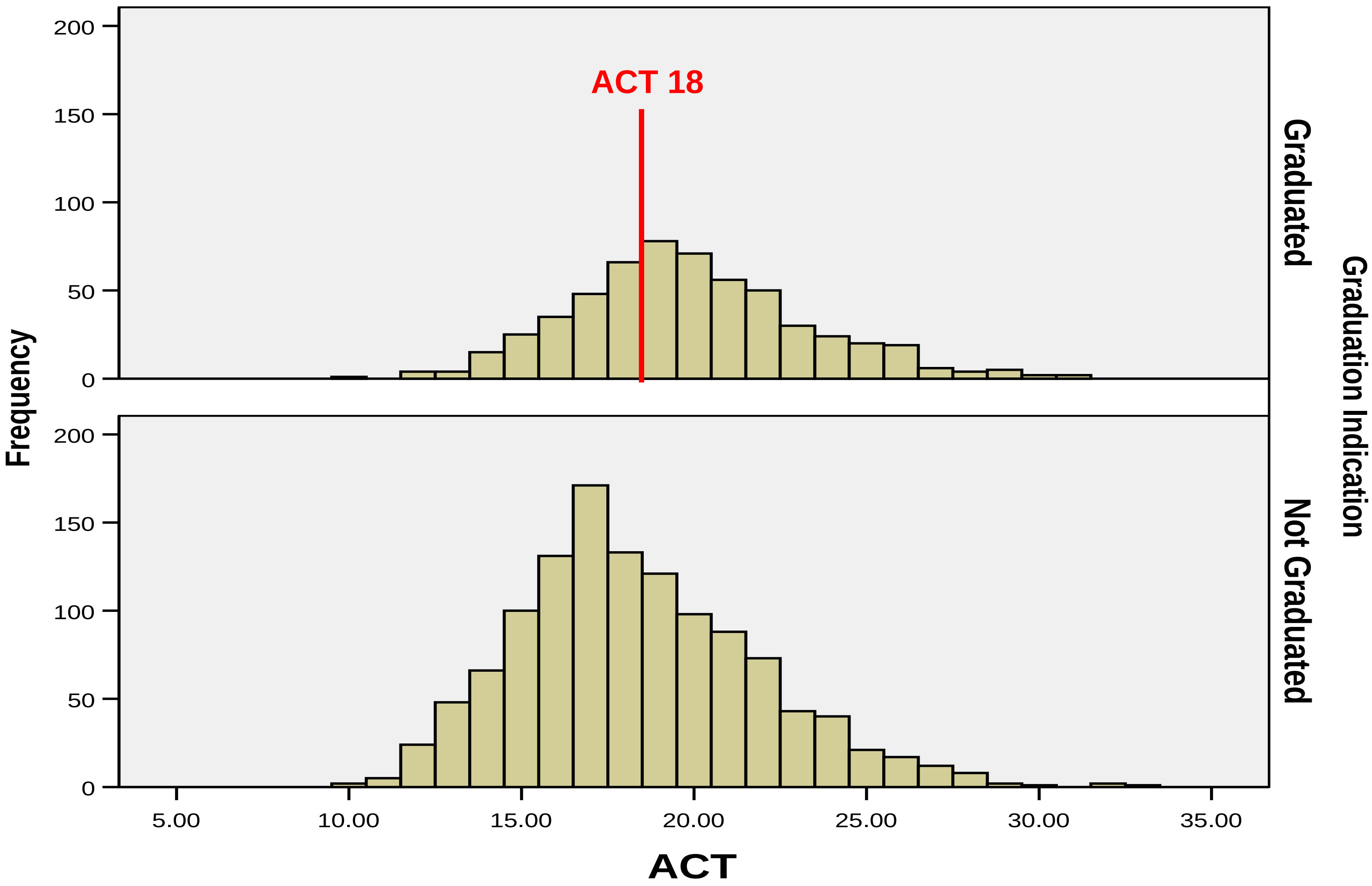
One year retention rate:

- Fall 2006 cohort → 70.0%

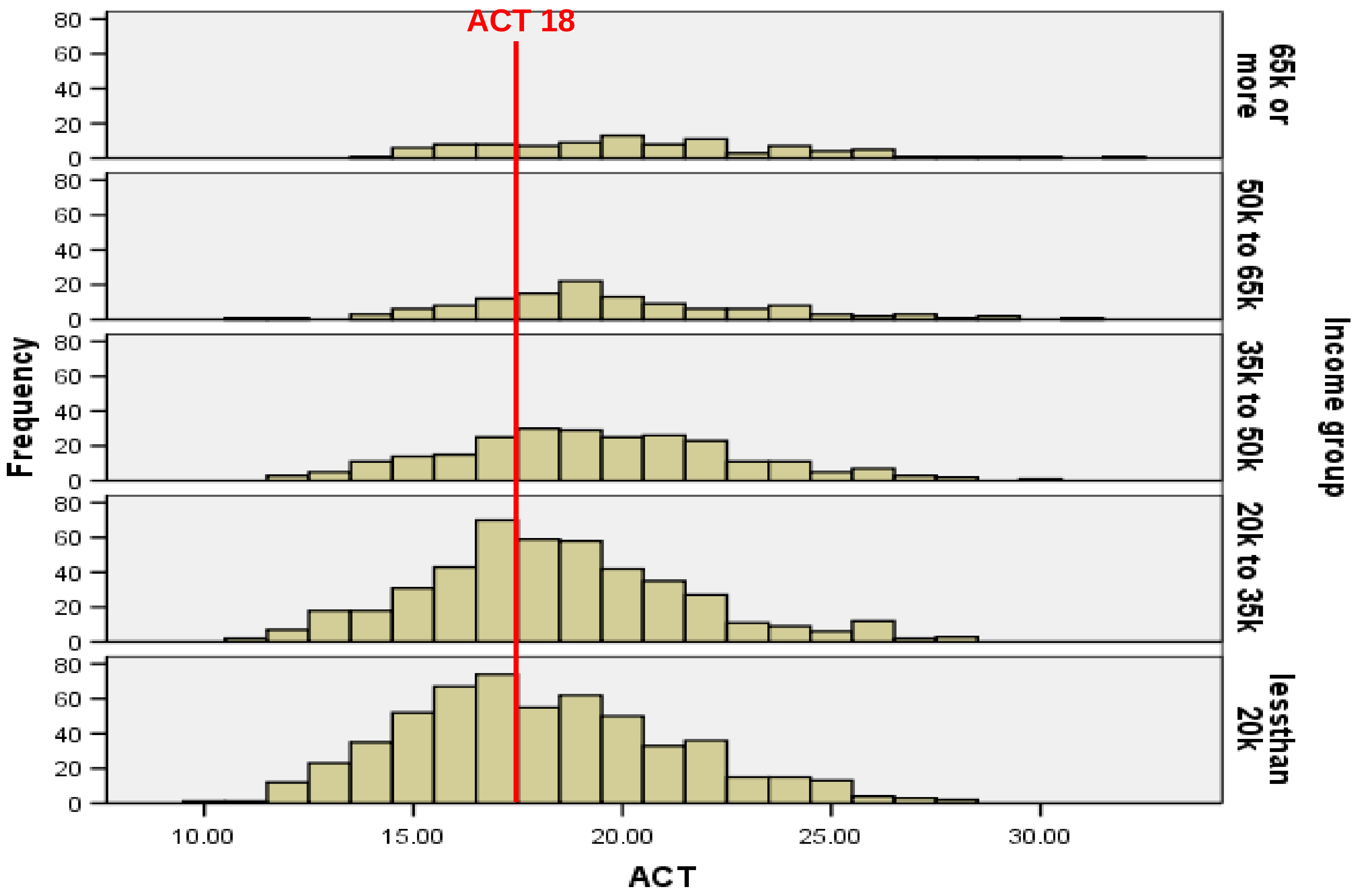
Practical Questions that Motivated this Project

- What more can the institution do to improve retention and degree completion rates?
- What more can the institution do to improve the effectiveness of our current initiatives?

The Range of Success



Impact of Selectivity



Main Research Questions

1. What factors affect student departure?
2. What factors impact students' return after an initial stopout?
3. What factors affect on-time graduation?
4. How can the institution identify (at the time of admission) students at-risk of departure?

Dataset

- 4,147 students who entered as first-time degree-seeking undergraduate freshmen during the 1999 and 2000 fall semesters*.

- The data set contained 12 semesters of fall and spring (six years) enrollments and non-enrollments arranged in a person-period format (where there are 12 records for each student, one for each semester).

*The size might differ from one model to the other due to case-wise deletion of missing values.

Research Questions 1 & 2

Factors That Explain Departure And Reenrollment

Methodology for Examining Factors that Predict Departure & Re-enrollment

Considerations:

- Student departure (dropout/stopout) and re-enrollment is a longitudinal process
- Critical to consider the *timing* of events and *factors that affect these events at each point in time*.

Approach:

- A *multi-spell discrete-time logit model* is used to analyze student departure and return separately.

Results: Factors That Explain Departure

Predictors of Departure:

- Low semester GPA will increase the risk of departure.
- Over time, the effect of semester GPA *decreases*.
- Failing a course will increase the risk of departure.
- Part-time enrollment increases the risk of departure.
- Financial aid (loans, grants, work study) will reduce the risk of departure.
- Students who stop-out are more likely to leave again.

Results: Factors That Explain Departure

- Timing Factors Involved in Departure
  - Risk of departure is higher in earlier semesters
  - Risk of departure is higher in spring semesters than fall semesters
- Upon Re-enrollment:
  - Risk of departure for returnees is higher in the semesters immediately after returning
  - Risk of departure for returnees increases with the length of the stop-out period

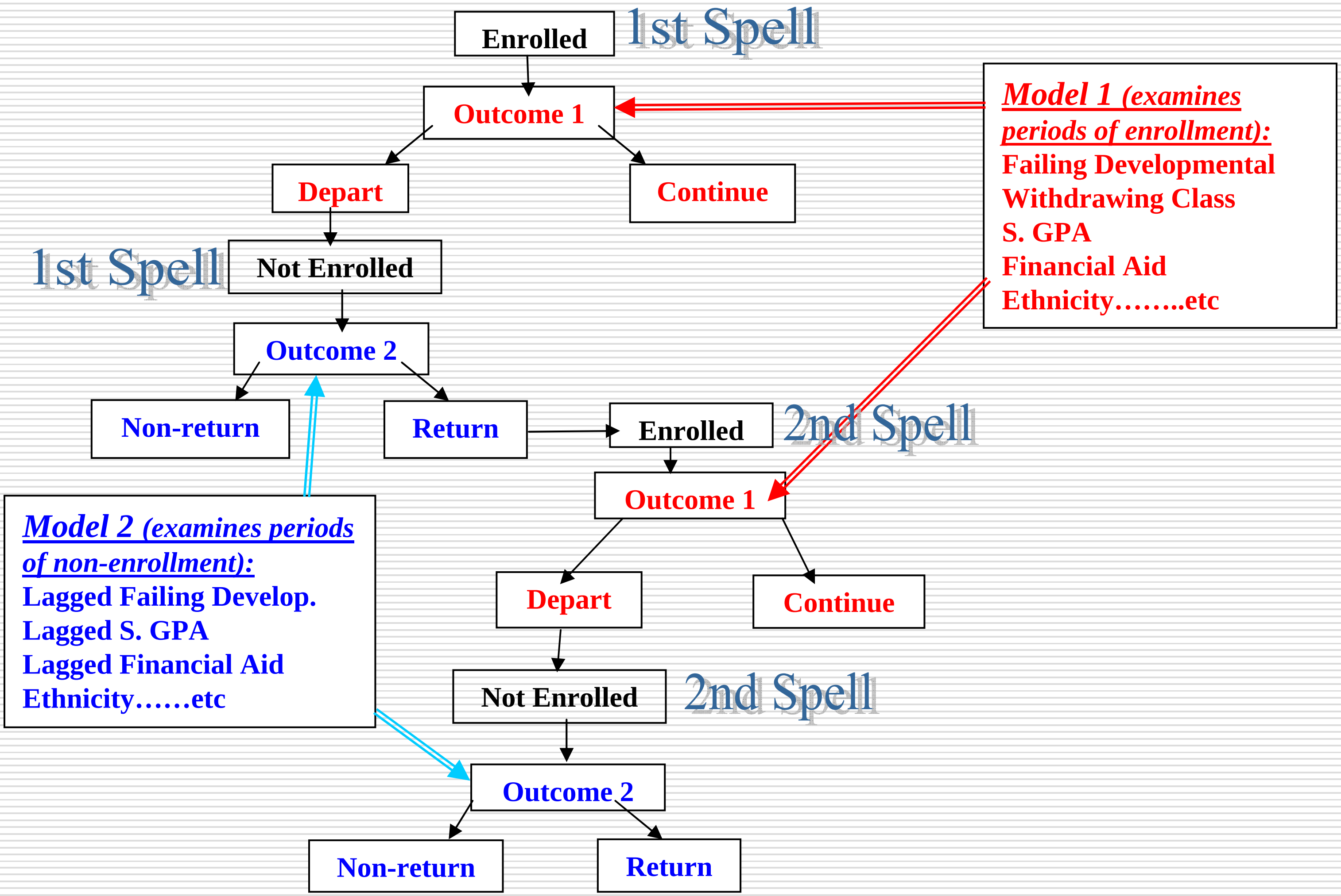
Results: Factors That Explain Reenrollment

The chances of returning to UTEP after stopping out:

- are higher for students with good academic standing at the time of departure
- are lower for older students (20 years or older when they first enrolled)
- increase as a student's initial length of enrollment increases
- decrease as a student's length of stopping out increases

Building the Multi-spell Discrete-time Logit Model

Departure and Return:
Conceptual Framework



Multi-spell Discrete-time Logit Model

- **Multi-spell:** refers to spells, or spans of time, when students were enrolled and/or not enrolled (that is, stopped out).
- **Discrete time:** refers to our use of semesters as discrete units, or measures, of time for examining the timing of events. Due to the nature of the academic year, the discrete measure of time in terms of semesters was appropriate.
- **Logit model:** incorporates a logit-link function (inverse of the logistic function) that relates an outcome with relevant predictor variables.

Departure and Return: Logit Model

The population discrete-time hazard for student *i* in semester *j* is (Willett and Singer, 2003):

$$h(t_{ij}) = \Pr\left[T_i = j \mid T_i \geq j \wedge X_{1ij} = x_{1ij}, \cdots, X_{Pij} = x_{Pij}\right]$$

The hazard is modeled using a logit link function and captures

- The general shape of the hazard profile (baseline hazard)
- The heterogeneity of the hazard caused by different predictor variables.

$$\log it[h(t_{jk})] = [baseline\_hazard\_profile] + [\beta_1 X_{1jk} + \beta_2 X_{2jk} + \cdots + \beta_P X_{Pjk}]$$

Departure: Empirical Results

Effect of timing variables:

Covariate	Coefficients (SE)
Intercept	-1.08(.07)***
year	-.54(.06)***
$year^2$	.04(.01)***
Returnee	1.86(.07)***
ln(semesters in spell)*Returnee	-.78(.07)***
ln(stopout length)* Returnee	.17(.06)**
Spring	.44(.07)***
Spring*year	-.09(.02)***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Departure: Empirical Results

Effect of academic performance variables:

Covariate	Coefficients (SE)
Part-Time Enrollment	.65(.05)***
Semester GPA (2.0+)	-2.26(.07)***
Semester GPA*ln(t)	.35(.05)***
Developmental	-.59(.10)***
Developmental*ln(t)	.73(.09)***
Failing Developmental	.72(.07)***
Withdraw one class	.49(.05)***
Withdraw two or more classes	.92(.09)***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Return: Empirical Results

Effect of timing variables:

Covariate	Coefficients (SE)
Intercept	-2.18(.07)***
ln(t)	-1.12(.04)***
Repeated Spell	-.26(.07)***
Fall	.54(.06)***
ln(Enrollment Spell length)	.51(.04)***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Return: Empirical Results

Effect of lagged variables:

Covariate	Coefficients (SE)
Lagged Part-Time	-.06(.06)
Lagged S. GPA	1.29(.14)***
Lagged S. GPA*ln(t)	-.42(.08)***
Lagged S. GPA*ln(spell one length)	-.47(.09)***
Lagged Failing Developmental	-.08(.06)
Lagged Withdraw one class	.03(.07)
Lagged Withdraw two or more classes	-.15(.09)

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Research Question 3

Predictors of Graduation

Phase I: Factors That Explain Graduation

- Stepwise Logistic regression model
- Variables explored for a significant effect on graduation:
  - Demographics
  - Academic preparation
  - Student perception (New student survey)
  - First term Academic Performance
- 21 variables were used in the model.

Phase I: Factors That Explain Graduation

- Significant factors:
  - Failing class in the first term will decrease the probability of graduation.
  - Higher first term GPA will increase the probability of graduation
  - Working 30 or more hours per week will decrease the probability of graduation
  - Students from bottom half of their high school class rank are less likely to graduate
  - Placing in below college level Math will decrease the probability of graduation

Limitations of the Logistic Regression

The model did not consider, for instance, the effect of:

- GPA over time (cumulative GPA)
- # of Semesters enrolled Part-time
- # of Developmental courses Enrolled (& Failed)
- Financial aid over time
- Stopping out

Phase II: Longitudinal Approach

- We used a longitudinal approach to explore the effect of stopping out as well as time varying effect of academic performance and financial aid.

- A proportional sub-distribution hazards model (class of survival model) was used to model student graduation under a competing risk setting.

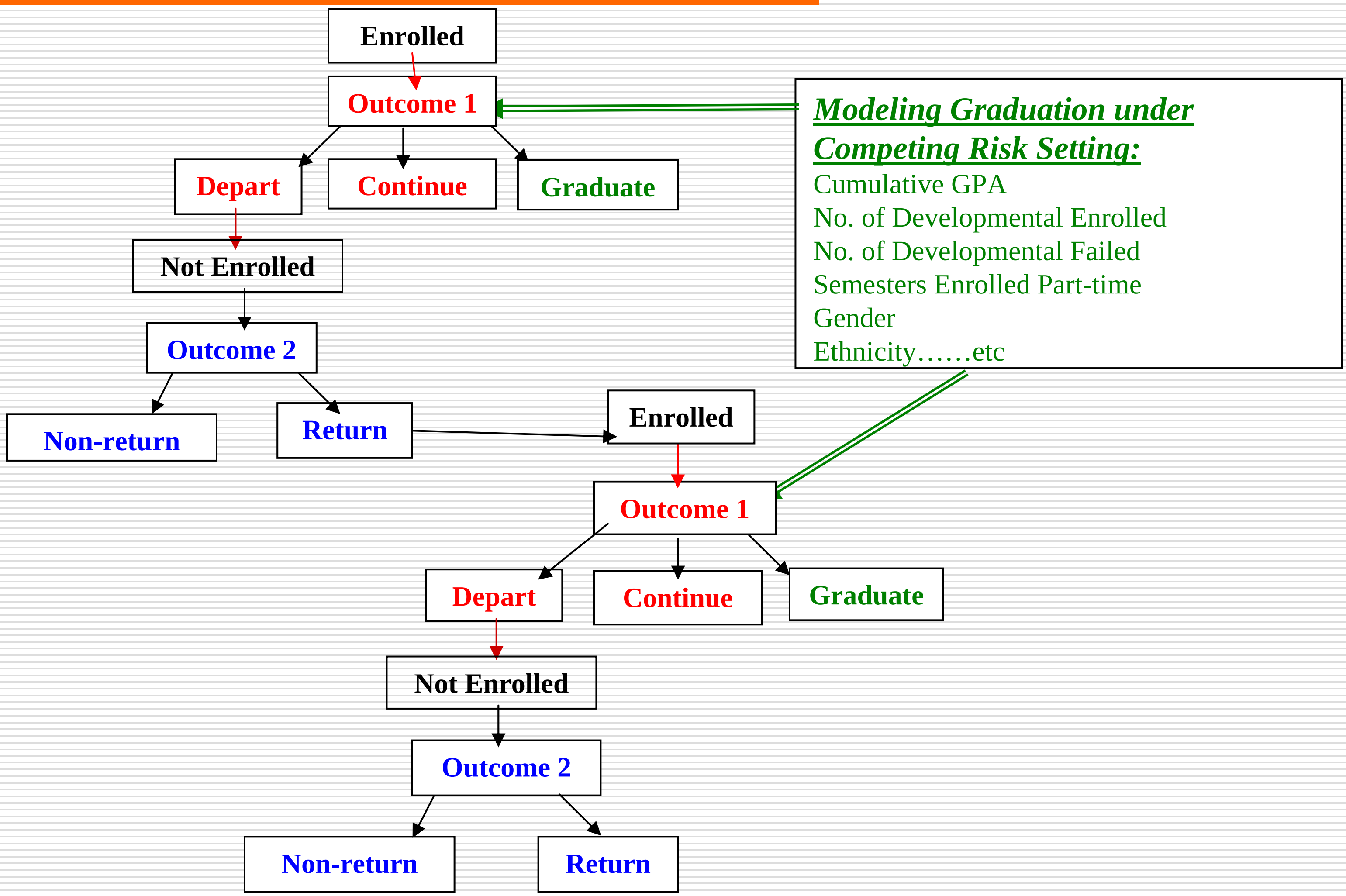
Factors That Explain Graduation

Competing risk model results:

- Higher cumulative GPA increases the probability of graduation
- Receiving financial aid (loans, grants and work study) increases the probability of graduation
- Failing any class decreases the probability of graduation
- Enrolling and passing a developmental course did not decrease a student’s chance of graduation significantly
- Stopping-out decreases the probability of graduation

Modeling On-time Graduation

Graduation: Conceptual Framework



Graduation: Competing Risk Model

- DesJardins et al. (2005) used a competing risk model to jointly model student departure and graduation.
- This study tries to model graduation under a competing risk setting without simultaneously estimating the departure model using the ‘Proportional Sub-distribution Hazards’ regression model described in Fine and Gray (1999).
- The model directly assesses the effect of covariates on the sub-distribution of a particular type of event (graduation) in a competing risks setting.

Graduation: Competing Risk Model

Our interest is modeling the cumulative incidence function for cause 1 (graduation) conditional on the covariates:

$$F_1(t, Z) = \Pr(T \le t, \varepsilon = 1 \mid Z)$$

Independent Variables

To formalize the subdistribution hazard, Gray's definition is:

$$\lambda_1(t; Z) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\{t \le T \le t + \Delta t, \varepsilon = 1 \mid T \ge t \cup (T \le t \cap \varepsilon \ne 1), Z\}$$

$$= -d \log\{1 - F_1(t; Z)\} / dt.$$

Under a proportional hazards specification with

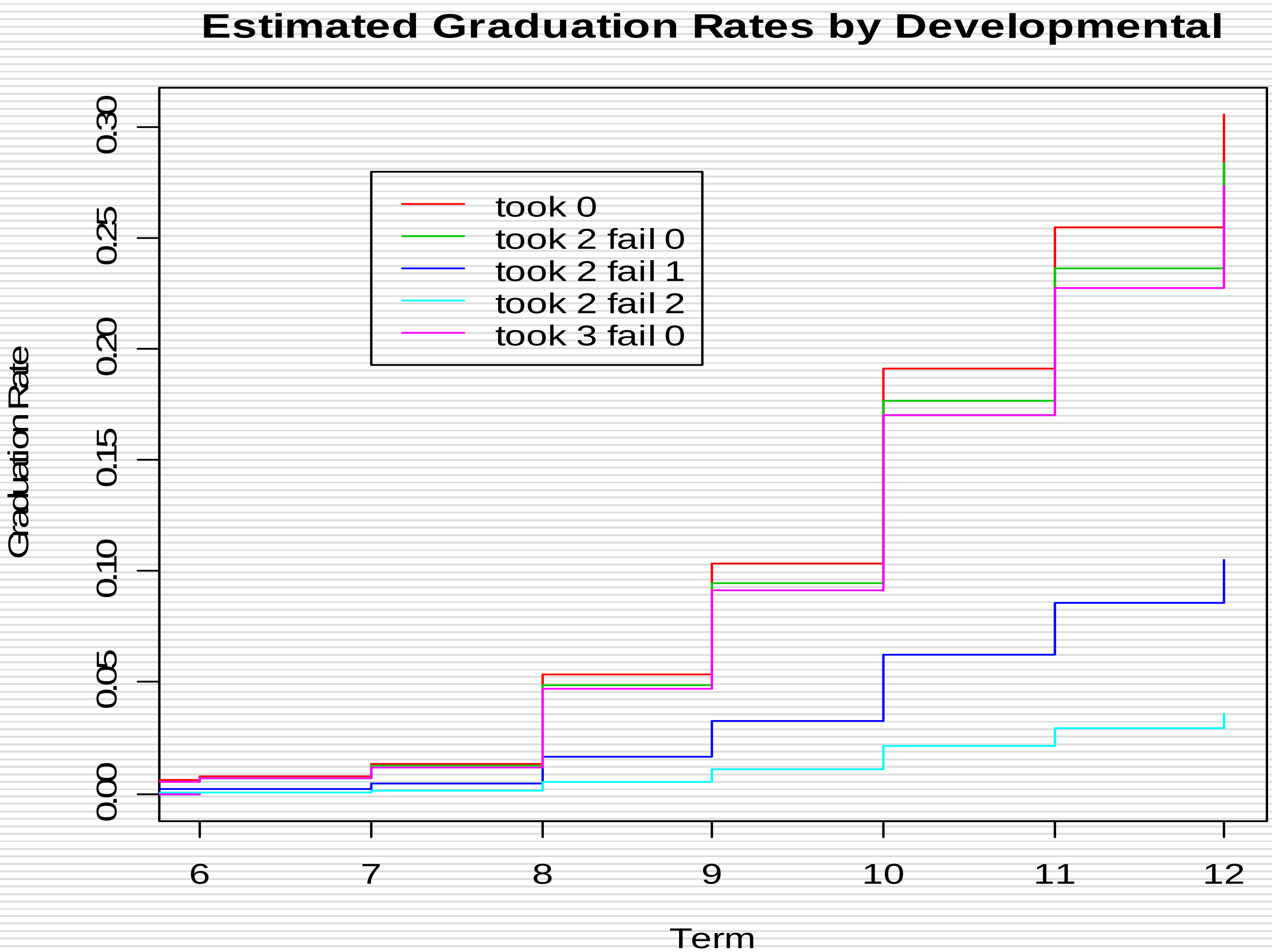
$$\lambda_1(t; Z) = \lambda_{10}(t) \exp(Z^T \beta_0)$$

Graduation: Empirical Results

Covariates	Coefficients (Standard Error)
Female	.06(.07)
Black	.25(.29)
White	.07(.13)
Asian	.04(.32)
International	.86(.11)***
Age20 or Older	-.54(.23)*
No. of Part-time Semesters enrolled	-.02(.03)
Cumulative GPA	1.58(.06)***
No. of Developmental courses enrolled	-.04(.03)
No. of Failed Developmental courses	-.28(.08)***
No. of Withdrawals	-.11(.02)***
No. of terms received Grants	.12(.01)***
No. of terms received Loans	.06(.01)***
No. of terms received Work-Study	.02(.02)

Significance code: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Graduation: Empirical Results



Research Question 4

Identifying Students
At Risk of Departure

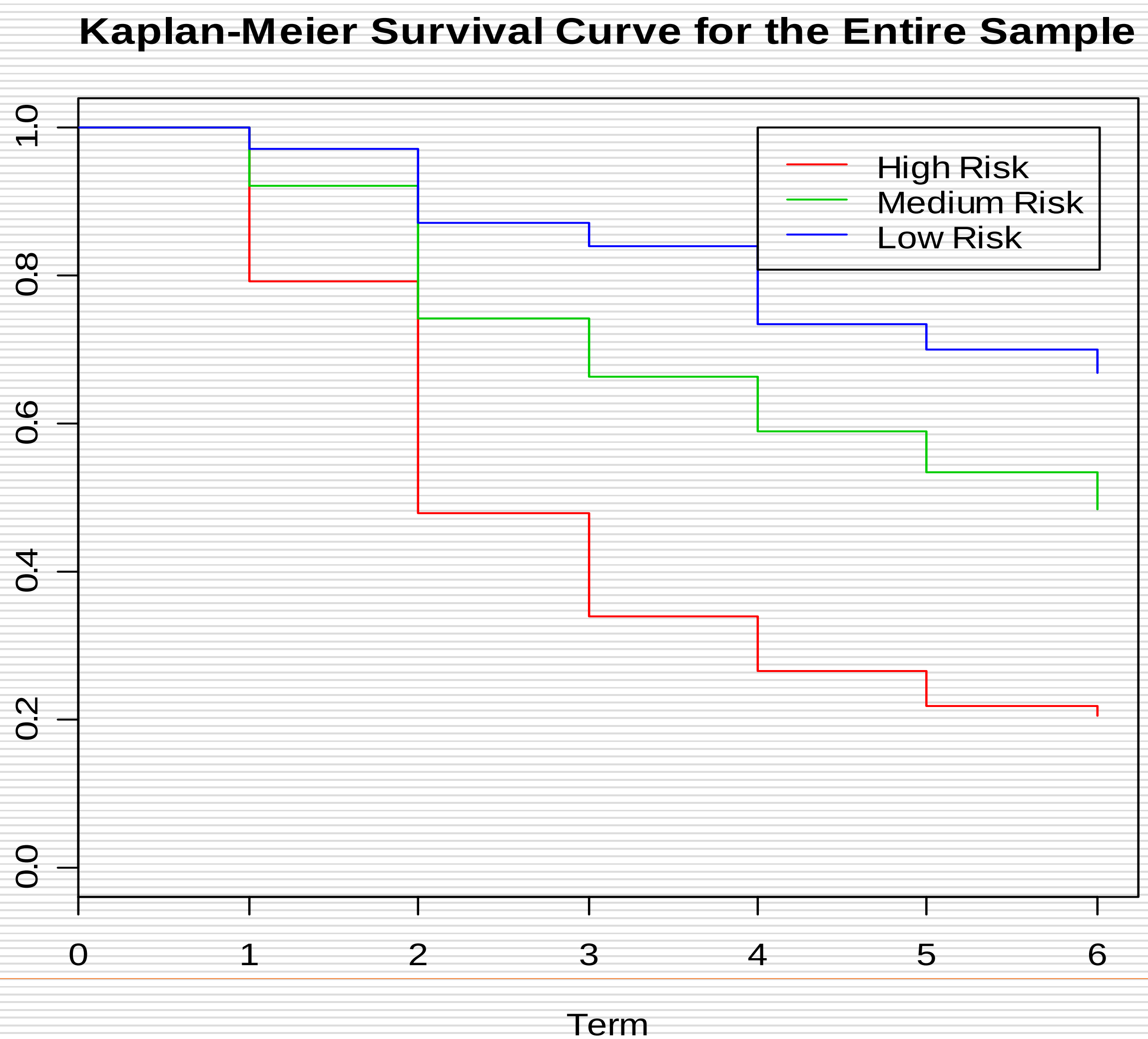
Identifying Students At Risk of Departure

- One way of improving retention rates and promote student success is to identify those students at risk of leaving and implement an appropriate **early** intervention according to the level of risk of the students.
- Cox proportional hazards (PH) regression model is used to determine the risk of departure for each student
- Significant factors at the <u>time of admission</u> that are used to identify level of risk:
  - Mathematics placement score
  - High school class rank percentile
  - Intended number of hours spent to work
  - Delayed matriculation from high school

Assigning Risk Score

- <u>Low risk</u> group: Students with 0 risk score
    - with college level math placement
    - top quartile of high school class
    - intend to work less than 20 hours per week
    - Directly matriculated from high school
- <u>Medium risk</u> group: Students with a score of less than 1.15 are assigned into medium risk group
- <u>High risk</u> group: Students with a score of 1.15 or higher are assigned into high risk group

Survival Trends Associated with Risk Groups



Retention Trends for Risk Groups

- One year retention rates:
  - High Risk: 48%,
  - Medium Risk: 74%
  - Low Risk: 87%
- Three year retention rates:
  - High Risk: 20%
  - Medium Risk: 48%
  - Low Risk 67%

Graduation Trends for Risk Groups

- Six year graduation rates:
  - High Risk: 9%
  - Medium Risk: 30%
  - Low Risk: 60%
- Graduation rate upon persisting through the first year with cumulative GPA of at least 2.0
  - High Risk: 25%
  - Medium Risk: 43%
  - Low Risk: 70%

Process for Generating Risk Scores

Cox Proportional Hazards Model

A Cox Proportional Hazards (PH) model is a (semi-parametric) class of survival models that relates the hazard rate, *h(t)*, with a set of predictor variables.

$$h(t) = h_0(t)\exp(\underbrace{\beta' x}_{\text{RS}})$$

Where

$h_0(t)$ is the baseline hazard function. It is assumed to be unknown and is left unparameterized.

$\beta' x$ are the predictor variables and regression coefficients. It is also known as the **Risk Score (RS)** and can be used to stratify students in to different risk groups.

Generating Risk Score

$$RS = \beta_1 x_1 + \cdots + \beta_p x_p$$

Validating the model

- The data was first divided randomly (in the ratio 3:1) into a training data set and a validation data set.
- Using the derived model the risk scores was calculated for each student in the validation set.

For further analysis

- It is also possible to estimate the parameters by bootstrap sampling where we sample 75% of the data and estimate the Cox model. This sampling and estimating the Cox model can be repeated many times; and the average of the estimated parameters can be used to generate risk scores.

Cox Regression Result

The Cox regression model and statistical test of the PH assumption

Cox Model				Checking PH Assumption		
Covariate	Coef.	SE (Coef.)	P	Rho	Chi. Sq.	P
BANM Level 2	0.216	0.1	0.031	0.011	0.135	0.714
BANM Level 1	0.441	0.09	0	0.005	0.031	0.86
HSP 2^{nd} Quartile	0.5	0.08	0	-0.036	1.431	0.232
HSP Bottom Half	1.697	0.09	0	0.028	1.006	0.316
Work 20 to 30 hpw	0.307	0.07	0	-0.034	1.247	0.264
Work more than 30 hpw	0.596	0.09	0	-0.031	1.109	0.292
Delayed Matriculation	0.247	0.13	0.06	-0.042	2.026	0.155
HSP Bottom Half*$time^3$	-0.003	0	0	0.194	78.12	0
N	1521					

β

Summary

- The major findings of this study shows the significant effect of:
  - Timing on departure and re-enrollment
  - (Length of) Stopping-out
  - Academic performance over time
  - Financial aid
- The study also:
  - Reveals the stratification of the student population into three risk groups with distinctive survival patterns

Implications and Further Research

- Implications
  - Evaluate efficacy of current interventions for each risk group, and modify interventions
  - Develop tools and methods to track progress during critical periods
- Areas for further research
  - Advance understanding of students within each risk group
  - Explore how ecosystems affect student performance
  - Explore factors that explain the success of transfer students

Changing the Face of Higher Education

"We are here to provide opportunities to a group of people who represent America's future. And it's such an awesome responsibility. And it's also an enormous satisfaction when you go to graduation and you see in front of you this group of graduates whose lives you know are going to be completely changed as a result of what's happened here. That's enormously exciting."

--Diana Natalicio
President of UTEP
Interview with Lumina Foundation for Education

References

Box-Steffensmeier J. M. and Jones B. S. (2004). *Event History Modeling: A Guide for Social Scientists*, Cambridge University Press.

DesJardins, S. L., Ahlburg, D. A. and McCall, B. P. (2002). *A Temporal Investigation of Factors Related to Timely Degree Completion*. The Journal of Higher Education, Vol. 73, No. 5 pp. 555-581.

DesJardins, S. L., Ahlburg, D. A. and McCall, B. P. (2006). *The Effects of Interrupted Enrollment on Graduation from College: Racial*, *Income, and Ability Differences*. Economics of Education Review, Vol. 25, Issue 6. pp. 575-590.

Jason P. Fine and Robert J. Gray (1999). *A Proportional Hazards Model for the Subdistribution of a Competing Risk*. Journal of the American Statistical Association, Vol. 94, No. 446, pp. 496-509.

References

Johnson, Iryna Y. (2006). *Analysis of Stopout Behavior at a Public Research University: the multi-Spell Discrete-Time Approach*. Research in Higher Education, Vol. 47, No. 8. pp. 905-934.

Packham C., Gray D., Silcocks P. and Hampton J. (2000). *Identifying the risk of death following hospital discharge in patients admitted with a suspected acute myocardial infraction in whom the diagnosis is not confirmed*, European Heart Journal, Vol. 21, Issue 3. pp. 213-219.

Singer, J. D., and Willet, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*, Oxford University Press, Oxford, New York.

Willet, J. B., and Singer, J. D. (1991). *From Whether to When: New Methods For Studying Student Dropout and Teacher Attrition*. Review of Educational Research, Vol. 61, No. 4, (Winter 1991), pp. 407-450.

Contact Information

Roy Mathew, Ph.D.
Director
Center for Institutional Evaluation, Research and Planning
The University of Texas at El Paso
Administration Building, Room 318
(915) 747-5117
rmathew@utep.edu

Denise Carrejo, Ph.D.
Assistant Director
Center for Institutional Evaluation, Research and Planning
The University of Texas at El Paso
Administration Building, Room 318
(915) 747-5117
dcarrejo2@utep.edu

Bereket T. Weldeslassie
Research Assistant
Center for Institutional Evaluation, Research and Planning
The University of Texas at El Paso
Administration Building, Room 318
(915) 747-5117
btweldeslassie@utep.edu

This AIR 2008 Presentation is available at:

http://cierp.utep.edu