



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Entrance Test Scores and College Dropout:
A Censored Quantile Regression Approach

Gustavo Ramos Sampaio
Department of Economics
Universidade Federal de Pernambuco
gustavorsampaio@gmail.com

Mary Arends-Kuenning
Department of Agricultural and Consumer Economics
University of Illinois at Urbana-Champaign
marends@illinois.edu

Breno Sampaio
Department of Economics
Universidade Federal de Pernambuco
brenorsampaio@gmail.com

Selected Poster prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.

Copyright 2013 by Gustavo Ramos Sampaio, Mary Arends-Kuenning and Breno Sampaio. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Entrance Test Scores and College Dropout: A Censored Quantile Regression Approach

Gustavo Ramos Sampaio¹, Mary Arends-Kuenning², Breno Sampaio¹

¹Department of Economics, Universidade Federal de Pernambuco, ²Department of Ag. Economics, University of Illinois at Urbana-Champaign



Objective

Our main objective in this paper is to analyze the relationship between entrance exam scores in Brazilian universities and their predictive power on the probability of college dropout.

Introduction

High university dropout rates have become a major concern in many countries and have attracted researchers' attention over the last decades (see, for example, Tinto (1993); W. Norton (1989); Oosterbeek (1989) and Hartog, Pfann, and Ridder (1989)). Many of these papers are devoted to analyzing the determinants of university dropouts, with specific attention given to the estimation of the relationship between entrance test scores and the probability of dropout. Obtaining correct estimates for the relationship between these two events is not as simple as it may seem from a first view. For the US, which has been the most analyzed country in the world, many complications arise from the fact that not only the SAT score determines acceptance in college education.

For Brazil, the situation is quite different. The admissions process and the institutional set up of the university makes the identification of the parameter of interest much more reliable. Three main advantages are responsible for that. First, all students are required to choose their major before taking the exam and, once they have entered the university, they are not allowed to switch majors. Secondly, once the student is enrolled at the university, the curriculum is the same for all students, with very small flexibility allowed for those students who are close to graduation. The third important difference is that acceptance is completely determined by the entrance test score. Thus, we need not worry about any other selection bias arising from unobserved variables that affect acceptance.

Data and Institutional Background

The dataset used in this paper comes from students who enter the Universidade Federal de Pernambuco (UFPE), which is a major university that is located in the Northeast of Brazil. UFPE is a public university which charges no tuition fees and the only requirement for entering the universities' undergraduate programs is an entrance exam, the vestibular, that must be taken by all candidates. Final acceptance for higher education depends not only on the final score obtained by the student, but also on which major the student applied for.

The data used in the paper includes all students who entered the university in the years of 2002, 2003 and 2004. There are about 2,600 students accepted each year at the university, leading to a total of 7,706 students in the sample.

Table 1: Summary statistics

Description	μ_x	σ_x	Min	Max
Entrance score	5.925	0.958	2.988	9.125
Dropped early	0.315	0.464	0	1
Last term	9.521	3.454	1	19
Age	19.975	4.113	15	55
First year	2003.019	0.804	2002	2004
Major			1	31
Mother education				
Elementary or less	0.077	0.266	0	1
Middle school	0.137	0.344	0	1
High school	0.380	0.485	0	1
College	0.404	0.490	0	1
Income				
<5 MWs	0.259	0.438	0	1
5-10 MWs	0.362	0.480	0	1
11-15 MWs	0.152	0.359	0	1
16-20 MWs	0.095	0.293	0	1
>20 MWs	0.130	0.336	0	1
Vestibular	1.193	1.126	0	4
N. of observ.	7,706			

Note: ¹The Brazilian national minimum wage is adjusted annually. The minimum wage in 2004 was R\$ 260,00 per month (which corresponds to R\$8,67 per day and R\$1,18 per hour).

Methods

Cox Proportional Hazard Model

We start by estimating a Cox proportional hazard model (see Cox (1972)) to analyze how entrance test scores and the decision to dropout from college are related

$$\lambda(t|ETS_i, x_i) = \lambda_0(t) \cdot \exp(ETS_i^T \beta + x_i^T \delta) \quad (1)$$

The model makes no assumption whatsoever about the form of $\lambda_0(t)$, allowing it be arbitrary, but assumes a parametric form for the effect of the predictors on the dropout hazard. Also, since the baseline hazard function is independent of the vectors of covariates, the parameters of the model will only shift the hazard function up or down. Thus, covariates must either increase the hazard over the whole time scale, if the estimated coefficient is positive, or decrease it, if it is negative.

The downside of using Cox's proportional hazard model to analyze the effects of entrance performance on dropouts is that it does not allow us to observe how covariates vary in time. Since the baseline hazard function λ_0 is non-negative, it only allows for the estimates to be either entirely above the horizontal "effect equals zero" axis or entirely below it. The model forbids the possibility that treatments may increase the hazard for a time and then decrease it.

Censored Quantile Regression

Quantile regression methods were first introduced by Koenker and Bassett (1978) and offer a mechanism for estimating models for the conditional median function, and the full range of other conditional quantile functions. Thus, it provides a more complete statistical analysis on how the variable of interest varies among the different quantiles. In such a way, we can observe how exam performance affects the probability of dropping out of college across the time quantiles. This is an important effect to capture since the probability of dropout might vary significantly over a student's college lifetime.

$$\hat{\beta}(\tau) = \arg \min_{\beta} \sum_{i \notin K(\tau)} \rho_{\tau}(Y_i - x_i^T \beta) + \sum_{i \in K(\tau)} \left\{ w_i(\tau) \rho_{\tau}(C - x_i^T \beta) + (1 - w_i(\tau)) \rho_{\tau}(Y_{+\infty} - x_i^T \beta) \right\} \quad (2)$$

Results

Performance on the vestibular exam appears to be negatively related to the probability of dropout. This implies that students who enrolled with low entrance test scores dropout with a higher probability when compared to students that enrolled with high entrance scores. This result, however, compares students enrolled in different majors and may reflect only selection of the best students enrolled in more competitive majors that require high entrance test scores for admittance. To account for differences among courses, we add course fixed effects in column 3. By comparing only students enrolled in the same course, we obtain that the coefficient for ETS becomes positive, i.e., students who enrolled with high ETS are on average dropping more than students who enrolled with low ETS.

Table 2: Cox Proportional Hazard Model

	(1)	(2)	(3)	(4)
Entrance Score	-0.022 (0.021)	-0.085*** (0.022)	0.094*** (0.038)	0.064** (0.038)
exp(β)	0.978	0.918	1.098	1.066
Year Fixed Effects	No	Yes	Yes	Yes
Major Fixed Effects	No	No	Yes	Yes
Other controls	No	No	No	Yes
N. of observ.	7,706	7,706	7,706	7,706

Results - Cont.

The technique developed by Portnoy (2003) allows us to combine survival analysis techniques, where data censoring is present, to the quantile regression techniques developed by Koenker and Bassett (1978). Figure 1 presents the censored quantile regression results. To compare with the Cox proportional hazard model estimates, we plot the results presented in Table 2. The solid line with no confidence bands indicates the estimated conditional quantile "effects" implied by the estimated Cox model. The vertical axis plots the effects of exam performance on college dropout rates, where positive effects indicate that the time before college dropout is increased, and the horizontal axis depicts the different time quantiles. The estimates are presented by the solid line contained in the blue areas, where the latter indicate a 95% confidence interval.

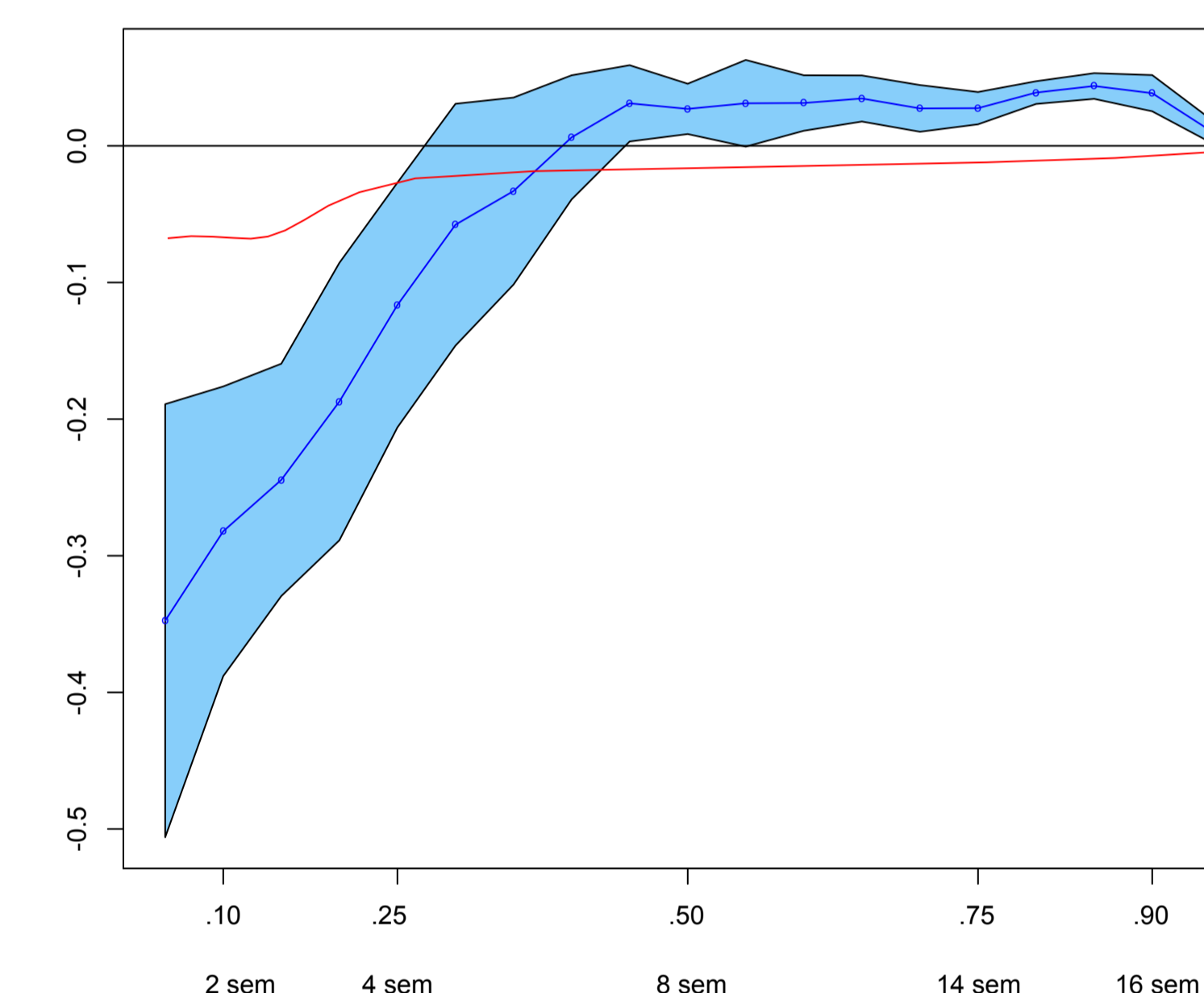


Figure 1: Conditional Quantile Effects on Time to Dropout

One can observe a better performance in the entrance exam increases the probability of college dropouts for the earlier quantiles (those below $\tau = 0.40$), but significantly decreases dropout probabilities for the latter quantiles. These results are consistent with past findings where positive correlations between entrance test scores and probability of college dropout are found. The intuition behind the observed high dropout probabilities following high entrance exam performance arises from the fact that high ability students, given the recognition of a wrong career choice, have a lower cost implicitly defined by their higher probability of being accepted for a different major on next years' exam.

Conclusions

The main findings is obtained via the Censored Quantile Regression Estimator where we find that a better performance in the entrance exam is correlated with a positive increase in the probability of college dropout for the earlier quantiles (those below $\tau = 0.40$), but significantly decreases dropout probabilities for the latter quantiles.

References

- Cox, D.R. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society. Series B (Methodological)* 34:187-220.
- Hartog, J., G. Pfann, and G. Ridder. 1989. "(Non-)graduation and the earnings function: An inquiry on self-selection." *European Economic Review* 33:1373-1395.
- Koenker, R., and J. Bassett, Gilbert. 1978. "Regression Quantiles." *Econometrica* 46:33-50.
- Oosterbeek, H. 1989. "An Economic Analysis of Educational Dropouts", University of Amsterdam. Paper presented at the European Economic Congress."
- Portnoy, S. 2003. "Censored Regression Quantiles." *Journal of the American Statistical Association* 98:1001-1012.
- Tinto, V. 1993. *Leaving college : rethinking the causes and cures of student attrition*. University of Chicago Press.
- W. Norton, G. 1989. "Dropouts, spells of time, and credits in postsecondary education: Evidence from longitudinal surveys." *Economics of Education Review* 8:49-67.