EXAMINING THE OUTCOMES OF THE EVALUATION SYSTEM FOR UNDERGRADUATE PROGRAMS IN BRAZIL: THE EFFECT OF SOCIOECONOMIC BACKGROUND, EDUCATIONAL CONDITIONS, SELECTIVITY, AND INSTITUTIONAL CHARACTERISTICS ON STUDENT PERFORMANCE

by

Cláudio de Albuquerque Marques

A Dissertation Submitted to the Faculty of the

CENTER FOR THE STUDY OF HIGHER EDUCATION

In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

In the Graduate College

THE UNIVERSITY OF ARIZONA

2003

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by CLÁUDIO DE ALBUQUERQUE MARQUES entitled EXAMINING THE OUTCOMES OF THE EVALUATION SYSTEM FOR UNDERGRADUATE PROGRAMS IN BRAZIL: THE EFFECT OF SOCIOECONOMIC BACKGROUND, EDUCATIONAL CONDITIONS, SELECTIVITY, AND INSTITUTIONAL CHARACTERISTICS ON STUDENT PERFORMANCE

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Dr. Dudley B. Woodard, Jr.		1/5/04 Date
Dr. Gary D. Rhoades		12/17/03 Date
Dr. John J. Cheslock		12/17/03 Date
Dr. Lee Sechrest		12/17/03 Date
		Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dr. Dudley B. Woodard, Jr.		1/5/04
Dissertation Director		Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under the rules of the Library.

Brief quotations from this dissertation are allowed without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED 

ACKNOWLEDGEMENTS

I wish to express my gratitude for those who helped me in so many different ways to make this project possible.

My special thanks to Dr. Woodard for inviting me into this program, for supporting me during my coursework, and for encouraging me during the dissertation process. Without his support and encouragement, I would not have reached this point.

Thanks to the faculty who supported me during my graduate studies as well as during this project. To Gary Rhoades, whose work and advice helped me develop a more critical view of higher education. To John Cheslock, for providing me the expertise and guidance necessary to develop this dissertation. To Lee Sechrest, whose enthusiasm for evaluation helped me develop a special interest for this field.

This project would not have been possible without the support of the Brazilian Government. First, my thanks to the *Universidade Federal do Ceará* and *Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior* for their support of my graduate studies. I would also like to express my appreciation for the *Instituto Nacional de Estudos e Pesquisas Educacionais* which made available the data used in this project.

My gratitude to the University of Arizona for offering me the financial support that made it possible to pursue my graduate studies. Special thanks to Roger Fox and Timothy Finan for taking care of the institutional agreement under which the support was provided. Thanks also to Roger Fox and Gail for sharing with my family many Thanksgiving and Independence Day.

My appreciation to Professor Roberto Claudio Bezerra for his encouragement to pursue my graduate studies abroad and for advising me to select the topic of my dissertation. To professor Tarquínio Prisco, for providing me very practical advice that help me move forward in my graduate studies.

I am thankful to Jesus Ramos whose support and friendship made my academic studies much more enjoyable than I ever expected. I will never forget your help, especially at the beginning of our program. Thanks also to José, Liang, George, Mayté, Lisa, Amy, Rich, Robyn, and Sylvia Mahon, among others, for their encouragement and support.

I wish to thank Christine and Heidi for their support in proofreading this dissertation.

My gratitude to some friends who supported my family and me. Special thanks to Aldevan, Ana, Kaina, and Wina, for sharing with us the most enjoyable moments that we had in Tucson. I wish to thank Aldevan, Robertinho and Careca for their friendship. Thanks to Jesus and Conchita, Guto and Ana Carvalho, Esteban and Fabiana, as well as their kids for their support that made our lives in Tucson so pleasant.

Finally, I wish to acknowledge the support received from my family. Thanks to my parents whose example, love, dedication and support helped me grow personally and intellectually. Many thanks to my father for his persistence in convincing me to pursue a graduate degree. Thanks to my mother and my mother-in-law for their prayers. Most importantly, I am grateful to my wife, my daughter, and my son who provided unconditional support that made this project a reality.

DEDICATION

I dedicate this dissertation to my wife, daughter, and son, for being the most important part of my life and whose love and support has been a blessing for my life and a source of inspiration.

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

TABLE OF CONTENTS – *Continued*

LIST OF TABLES

LIST OF TABLES – *Continued*

LIST OF FIGURES

ABSTRACT

This research project represents the first national effort to examine the evaluation system implemented in the Brazilian higher education system to assess undergraduate education and determine its usefulness as an information tool for public policymakers. Due to the significant postsecondary education enrollment growth since 1995, concern about the quality of undergraduate programs has dominated the agenda of higher education. With little research available to inform policymakers, no general principles exist to guide public policy formation.

A quantitative approach is used to explore the relationships between the quality of programs assessed by the program evaluation (ACO) and student performance at the exit exam (ENC). The data originated from three national databases in Brazil: the student exit exam (ENC) database, the evaluation of educational conditions (ACO) database, and the national information system for higher education (SIEdSup). The final sample comprises 77,085 undergraduate senior students from 698 undergraduate programs in four academic fields: Administration and Law (the largest fields), Civil Engineering (technological field), and Dentistry (a very selective field).

Findings from this investigation offer strong evidence to conclude that administrative control is a higher imperative than institutional type. Public institutions, especially federal and state institutions, provide education at a higher quality level than private institutions. In regard to the educational conditions, faculty is undoubtedly the most critical resource for success in college, although student learning can be enhanced

by having good quality curriculum and facilities. Lastly, SES background and program selectivity were significant predictors of student performance.

The findings from this research project lead to three general principles that are helpful in informing public policymakers. Investment in public education pays, as the educational outcomes of public institutions were higher than those from private institutions. Investment in faculty is required to promote high quality education, which in turn does not require high quality facilities. Investment in graduate education is extremely necessary in this period of rapid enrollment growth, particularly to provide institutions of higher education with more qualified faculty and reduce the existing gap between elite and non-elite institutions. Other implications and recommendations are also discussed.

CHAPTER I

INTRODUCTION

Overview

This dissertation examines the outcomes of the evaluation system put into practice in 1996 to evaluate Brazilian undergraduate programs. The term 'program' is used in this dissertation to refer to undergraduate programs (equivalent to four-year programs in the United States) offered by colleges and universities, including Law and medicine, which are not considered professional programs in Brazil.

This evaluation system was designed by the Brazilian federal government to provide information to higher education stakeholders and policymakers about the quality of undergraduate programs in a period of fast enrollment growth and increasing pressure for accountability. Two evaluation instruments were implemented by this system: the *Exame Nacional de Cursos* (ENC), which means national exit exam of programs, was designed to measure some skills, knowledge, and abilities developed by senior students during college (Maia Filho, Pilati, & Lira, 1998), and the *Avaliação das Condições de Oferta dos cursos de graduação* (ACO), which means evaluation of the offering conditions of undergraduate programs, was designed to assess the conditions under which education has been delivered in the following domains: faculty qualifications and working conditions, curriculum and pedagogy, and infrastructure (Ministério da Educação e do Desporto [MEC], 1999).

Undoubtedly, issues related to undergraduate program quality have been at the forefront of the agenda of the Brazilian Higher Education system since this evaluation system was implemented. In fact, the concern about institutional evaluation (or self-evaluation) was present mostly among universities at the time this evaluation system was proposed, due to a specific program known as PAIUB (institutional evaluation of Brazilian universities). This institutional evaluation was initiated in 1993 as a result of a joint effort between universities and the federal government (Ristoff, 1996). However, it was after the implementation of the first ENC in 1996 that concern about undergraduate program quality began to affect higher education institutions in a significant way (Dias Sobrinho, 1999; Maia Filho, Pilati, & Lira, 1998). Since then, the federal government has used the results of this evaluation system in the authorization and re-accreditation process of undergraduate programs (Ferrer, 2001; Gomes, 2002). The media has extensively publicized the outcomes of this evaluation system, mainly by ranking undergraduate programs and higher education institutions (Segenreich, 2000). Some colleges and universities have shaped institutional actions based on this evaluation system (Caldeira, Kraemer, & Vasconcelos, 2000; Freitas, 2000; Martins, 2000; Silva, 2001). Other researchers have discussed potential negative implications associated with the implementation of the ENC (Bastos, Michael, & Oliveira, 2001; Martins, 2000; Santos, 2001; Santos Filho, 1999).

Despite the enormous publicity and discussion raised since the implementation of this evaluation policy, issues crucial to understanding this evaluation system as a process capable of promoting significant changes in the higher education system have been

largely unexplored. What are the purposes, usefulness, and limitations of this evaluation system and its evaluation instruments? Do the current evaluation instruments gather and produce sufficient data to guide program improvement? What are the potential biases in the evaluation instruments that might threaten the interpretation and usefulness of evaluation outcomes? What are the relationships among the outcomes of this evaluation system? What are the short-term and long-term implications of the implementation of this evaluation policy with respect to the reorganization of the Brazilian Higher Education system?

More salient from the standpoint of policymakers is the fact that this evaluation policy has an enormous potential to reshape the Brazilian Higher Education system for better (or worse) by providing data and incentives to guide (or misguide) institutional actions. Some have argued that this policy has promoted a significant change in the relationship between the state and higher education institutions (Freitas, 2000; Gomes, 2002; Silva, 2001). Therefore, it is indispensable for the Brazilian Higher Education stakeholders and policymakers to have a comprehensive understanding of the design of this evaluation system, particularly with regard to how undergraduate program quality is conceptualized, measured, and monitored. In addition, it is helpful to identify the incentives embedded in this evaluation policy, as well as the limitations and biases that might threaten the utilization of the outcomes. Such an understanding would help promote the necessary improvements in this evaluation system.

This chapter is organized as follows. Initially, a brief discussion about the evaluation system for undergraduate programs is provided. Then, I state the problem that

led to the development of this investigation. Next, the purpose of the investigation is presented, followed by the research questions. Then, a brief discussion of the literature to be reviewed and the theoretical and conceptual frameworks are provided. Following that, the methodology selected to explore the research questions is presented. Finally, the scope, limitations, and organization of the study are presented.

Background

In the last decade, concern about the quality of undergraduate programs in Brazil has increased substantially as a result of rapid enrollment growth coupled with the implementation of a national evaluation system for undergraduate programs. The low access rate to higher education, when compared with other Latin American countries, has created pressure for the Brazilian Higher Education system to promote greater access (Durham, 1993). In order to achieve this goal, the federal government has encouraged enrollment expansion, especially in the private sector (Gomes, 2002; Silva, 2001).

The evidence that the federal government has successfully expanded access to higher education is presented in Table 1.1. From 1995 to 2001, total undergraduate enrollment increased by 72 percent, which represents an addition of 1.3 million new students enrolled. So far, private institutions have shown more flexibility in increasing their participation in higher education than public institutions, showing a clear trend to continue increasing their presence in Brazilian Higher Education. While enrollment in private colleges and universities increased by 97 percent in seven years, enrollment in public colleges and universities increased by 34 percent. This fact might support the

Table 1.1 - Undergraduate enrollment between 1995 and 2001

Year	Public			Private			Total	
	Enrollment		% Enrollment Growth since 1995	Enrollment		% Enrollment Growth since 1995	Enroll-ment	% Enrollment Growth since 1995
	Number	%		Number	%			
1995	700,540	40%	-	1,059,163	60%	-	1,759,703	-
1996	735,427	39%	5%	1,133,102	61%	7%	1,868,529	6%
1997	759,182	39%	8%	1,186,433	61%	12%	1,945,615	11%
1998	804,729	38%	15%	1,321,229	62%	25%	2,125,958	21%
1999	833,093	35%	19%	1,544,622	65%	46%	2,377,715	35%
2000	887,026	33%	27%	1,807,219	67%	71%	2,694,245	53%
2001	939,225	31%	34%	2,091,529	69%	97%	3,030,754	72%

Source: MEC/INEP/SEEC

claim that the federal government created an educational policy to promote more access to higher education through increasing the presence of private institutions in Brazilian Higher Education. In fact, the percentage of students enrolled in private institutions grew from 60 percent in 1995 to nearly 70 percent in 2001.

While enrollment growth promotes new educational opportunities for many students excluded from higher education, it might affect undergraduate program level quality if the enrollment expansion is not followed by institutional efforts designed to enhance quality. Sampaio (1998) discusses three assumptions that seem to be associated with the emergence of quality as an important issue in the agenda of Brazilian Higher Education. The first assumption is that public institutions provide education at higher standards of quality than private ones. This assumption is fundamentally based on the model of research universities that guided the expansion of the Brazilian Higher Education system in earlier decades. The second assumption is that there is a trade-off

between quality and enrollment growth, especially when the enrollment growth is promoted in a short period of time, as has happened in Brazil. Along these lines, the trade-off between educational quality and access in U.S. state universities has been investigated by Seneca and Taussig (1987). The third assumption is that in the absence of a quality-control policy or mechanism, higher education institutions will not take the necessary actions to promote high quality standards based on best practices (Sampaio, 1998). For that reason, Schwartzman (2003) supports government intervention in higher education, arguing that the Brazilian federal government has to ensure that colleges and universities are providing education at a certain standard of quality.

Based on these assumptions, the Brazilian federal government[1] implemented a national evaluation policy for undergraduate programs. The federal government acts as a self-regulator by providing the legal framework to which public and private higher education institutions must comply. According to Maia, Filho, Pilati, and Lira (1998), the goal of the evaluation policy is to guarantee that expansion of the system will occur by adhering to a minimum standard of quality. Gomes (2002), on the other hand, argues that this evaluation policy was implemented to regulate the expansion of the higher education system while increasing the control over the higher education institutions, creating a modern market for higher education. In order to achieve that, the evaluation system would make quality indicators available to higher education stakeholders through evaluation of programs related to student performance in an exit exam, as well as of

---

[1] The central government involvement in this quality assurance effort can be explained by the central role played by the Brazilian federal government in the higher education system, being at the same time provider, financer, and regulator of the system.

educational conditions (faculty, curriculum, and facilities). These indicators would be used by students to choose the college they attend, by the federal government to guide the authorization and accreditation processes, and by higher education institutions to guide institutional actions.

This evaluation system includes two components designed to evaluate undergraduate programs. The first component is a national exit exam (ENC), offered once a year since 1996. This exam was designed by the federal government to measure the skills, knowledge, and abilities developed by senior students as a result of attending college (Maia, Filho, Pilati, & Lira, 1998). The outcome of the ENC is a five-letter scale ranging from A to E that is determined based on the average performance of students on the exam. The expected outcome is an indicator that shows the relative position of a given undergraduate program compared to other programs in terms of the mean test score of the students.

The second component of this evaluation system is an external evaluation of undergraduate programs (ACO). External specialists visit each program once every five years. Visits are organized by a federal agency responsible for this external quality audit, which sets guidelines and provides enumeration of program dimensions to be audited, along with the documentation to be filled out before the visit. This evaluation was designed to assess the conditions under which education has been delivered in the following three dimensions: infrastructure, curriculum and pedagogy, and faculty qualification and working conditions. Programs visited and assessed by specialists receive an evaluation rating for each of these three dimensions, assigned from a four-

concept scale ("very good" condition, "good" provision, "regular" condition, and "insufficient" condition) according to the level of compliance with the established standards and guidelines. This scale represents the extent to which undergraduate programs have achieved or exceeded the specified quality standards defined by a commission of specialists in each of these dimensions.

The rationale behind this evaluation metric is that student performance at this national exit exam is assumed to be a function of the conditions of the undergraduate programs assessed by the external program evaluation (Silva, 2001). That is, students in programs with "very good" conditions in the dimensions evaluated by specialists would have many more educational opportunities to learn more, and in turn would also have higher performance in the exit exam. Consequently, high quality undergraduate programs would be those programs whose students performed well in this national achievement test. A further assumption is that exposing the quality level of undergraduate programs will lead higher education institutions to promote program improvement (Segenreich, 2000), which would be motivated by either the competition among institutions for public recognition and resources or the pressure exerted by the central government and the general population, especially students, for promoting better quality.

Statement of the Problem

Educational policies are designed to provide policymakers and stakeholders with the foundations necessary to promote educational changes. In the case of the evaluation system being examined in this investigation, the Brazilian federal government put into practice a national evaluation system whose objective is to assess the quality of

undergraduate programs. According to MEC (1999), "The evaluation process of programs aims to show institutions that took part at the national exam of programs [ENC], by specific recommendations, those aspects identified in the evaluation visits [ACO] as being subject of improvement" (p. 11). This assertion reveals an important assumption embedded in this evaluation policy. By implementing two evaluation procedures designed to assess the knowledge, skills, and abilities that students have developed during college (ENC), as well as specific dimensions of undergraduate programs (ACO), the federal government implies that student performance at the ENC is largely determined by some structural characteristics of the undergraduate programs assessed by ACO (Silva, 2001).

Despite the fact that the ENC and ACO were implemented in 1996 and 1997, respectively, the relationship among their outcomes has been largely unexplored by higher education scholars. The vast majority of scholars have mainly discussed the potential negative implications of the ENC, such as creating a national curriculum, disregarding local needs and institutional missions, and teaching students to take the exam (Martins, 2000; Santos, 2001; Santos Filho, 1999). Only a few researchers have addressed the methodological limitations of the ENC (Bastos, Michael, & Oliveira, 2001). Soares, Ribeiro, and Castro (2001) applied a value-added approach to explore some aspects that affect student performance on the ENC; however, they did not include the outcomes of the ACO in their analysis. No scholar has carried out an in-depth analysis of the relationship between the outcomes of ENC and ACO. More troublesome is the fact that there is no mechanism designed to analyze the convergence between the

results of the ENC and ACO (Santos, 2001). Even the data collected by the ENC and ACO are not yet integrated in a single database, which makes it very difficult to analyze the link between the outcomes of these evaluation processes.

The need for exploring this relationship can be justified by the fact that the ENC and ACO provide the rational ground for the incentives that inform policymakers and stakeholders about the actions institutions, governmental agencies, faculty, and students should take in order to comply with the notion of quality underlying this evaluation metric. In addition, the statement provided by MEC (1999), cited above, assumes that institutions would be able to help students to perform better in the ENC by promoting actions to improve those domains assessed by ACO. Indeed, some researchers point out that institutions of higher education have taken actions motivated by the logic of this evaluation system, aiming to achieve better outcomes in this evaluation system (Caldeira, Kraemer, & Vasconcelos, 2000; Freitas, 2000; Martins, 2000; Silva, 2001). However, research on higher education has not explored the relationship between the most important outcomes of this evaluation system, which is necessary in order to examine whether the incentives established by this evaluation system would promote better quality undergraduate programs.

Purpose of the Study

The purpose of this investigation is to examine the relationship between the evaluation outcomes of the ENC and ACO. Specifically, I want to examine the degree to which student performance at the ENC is related to program characteristics assessed by ACO. This analysis is primarily guided by research (Astin, 1991; Frazer, 1992; Willms,

1992) suggesting that educational outputs are supposed to be influenced by educational inputs, which in this investigation are represented by program characteristics assessed by ACO (faculty, curriculum, classrooms, laboratories, and libraries). In addition to the ACO evaluation outcomes, this investigation also attempts to explore the degree to which academic achievement in college is influenced by institutional characteristics as well as individual characteristics, as it has been shown in the literature on U.S. higher education (e.g., Pascarella & Terenzini, 1991). Therefore, I want to provide higher education stakeholders and policymakers with interesting insights about the relationship between the outcomes of the ENC and ACO, presenting and exploring possible reasons why some of the relationships mattered and why others did not seem to matter. As I bring the theoretical lens I have selected to explore the findings, the main focus is on making meaning of the results, as well as on defining the context necessary to understand the relationship between the ENC and ACO evaluation outcomes.

Research Questions

In order to accomplish the stated purpose, this investigation pursues two major research questions:

What are the relationships, if any, among the mean test score (ENC), faculty evaluation rating, curriculum evaluation rating, facility evaluation rating, program selectivity, and institutional characteristics (institutional type and administrative control)?

This research question investigates the extent to which the mean test score (ENC) is related to the outcomes of the ACO assessment at the program level. Its significance

can be explained by the fact that the evaluation outcomes are consolidated by undergraduate programs, aiming to provide information for accountability, monitoring undergraduate program performance, and inferring program quality. Nevertheless, no empirical evidence exists to show how these outcomes relate to each other. Due to significant institutional diversity in the Brazilian Higher Education system, I also explore the degree to which the relationships addressed in this research question change across institutional type and administrative control.

To what extent is student test score on the ENC associated with faculty evaluation rating, curriculum evaluation rating, and facility evaluation rating, while holding constant SES background, individual characteristics (race and age), program selectivity, and institutional characteristics (institutional type and administrative control)?

This research question examines the main assumption underlying the national evaluation policy of undergraduate programs; that is, student performance on the ENC is influenced by the quality of faculty, curriculum, and infrastructure assessed by ACO. In order to explore other dimensions that are likely to affect the association of student test score and educational conditions assessed by ACO, the influence of socio-economic status (SES) background, individual characteristics, institutional characteristics, and program selectivity are also explored in this research question.

Theoretical Frameworks

This investigation is tied to three conceptual frameworks briefly presented in this section: models for monitoring outcomes in higher education, institutional stratification theory, and the concept of cultural capital.

Models for Monitoring Outcomes in Higher Education

Frazer (1992) argues that any process designed to assess quality of teaching in higher education should include a comprehensive approach involving inputs, processes, and outputs. Following the same line of thought is work by Astin (1991), Bosker and Scheerens (1994), and Willms (1992). This work offers some models designed to monitor educational outcomes that examine the effectiveness of programs with regard to the transformation of educational inputs into educational outputs.

Willms (1992) offers a multi-level model (school, teachers, and pupils) designed to monitor student performance in schools. Although the input-process-output model addresses student performance in schools, some concepts and dimensions of this model are useful in guiding the investigation of student performance and program performance in Brazilian colleges and universities. The application of the input-process-output (IPO) model to higher education is based on the assumption that student performance is influenced not only by input factors (student characteristics, faculty characteristics, instructional resources, curriculum characteristics, class size, college size, per-student expenditure), but also by process factors (student's attitudes toward college, working conditions, student interaction, faculty orientation).

What is missing from Willms' model is the possibility that contextual aspects such as institutional selectivity and human and financial resources affect not only the college inputs but also the educational process, which in turn affects college outputs. In order to overcome this limitation, Bosker and Scheerens (1994) provide tools for considering some college characteristics as being contextual factors affecting both inputs

and processes in Willms' model. Bosker and Scheerens argue that school effectiveness is highly influenced by selection mechanisms that affect the quality of students, teachers, and administrators, a thesis that has been shown to also work in colleges and universities. Therefore, the inclusion of contextual aspects at the college level makes it possible to investigate the degree to which type of administrative control (public vs. private), institutional selectivity, socio-economic status (SES) of the student body, and qualifications of the faculty body affect student performance.

Astin (1991), in his book "Assessment for Excellence: the Philosophy and Practice of Assessment and Evaluation in Higher Education," provides the conceptual ground for criticizing the rational assumptions, as well as the educational outcome, used in the Brazilian evaluation system for undergraduate programs. Although this work presents a model for analyzing the educational effectiveness named input-environment-outcome model (I-E-O), he clearly opposes the use of standardized tests as a valid outcome for higher education institutions, as has been done in the Brazilian case. His point of view is based on the talent development approach used to define quality of education, which argues that the main focus of an evaluation in higher education institutions should be the assessment of multiple talents rather than the application of standardized tests.

These works are useful for this investigation because they provide the framework for analyzing and criticizing the relationship between the outcomes of the ENC and ACO implemented in the Brazilian Higher Education system to assess the quality of undergraduate programs.

Institutional Stratification Theory

Institutional stratification theory (Trow, 1984) addresses the mechanisms by which higher education institutions are stratified through the influence of government or markets. Trow argues that governments have the power to exert an enormous influence over the organization of higher education systems and the determination of the institutional status by using specific legislation and state policies to define the boundaries that create different groups of institutions. On the other hand, the application of market principles to higher education is responsible for institutional differentiation within the sector of higher education by awarding institutions with reputation and prestige, which creates the competitive advantage in the marketplace. Although Sampaio (1998) does not make any connections with institutional stratification theory, she points out that the problem of low quality in Brazilian Higher Education is associated with the absence of educational policy to control quality along with the privatization of higher education under market forces.

The importance of developing a better understanding of the influence of institutional characteristics on educational outcomes has received some attention from scholars in education. Pascarella and Terenzini (1991) cite some institutional characteristics, such as student body selectivity, institutional prestige or reputation, institutional resources, and facilities (for example, library size) as being frequently used as indicators of quality in the higher education literature. Bosker and Scheerens (1994), when examining performance in a lower education level (k-12), argue for the inclusion of some school characteristics when analyzing school effectiveness because "school

effectiveness is largely determined by selection mechanisms (effective schools are schools that attract good students, good teachers and good administrators)" (p. 168).

This claim also seems to work for colleges and universities. Elite institutions tend to attract the best resources available for higher education due to good reputation, financial resources, qualified faculty, highly selective admission processes, and a good infrastructure. This situation is likely to create a virtuous cycle for the most prestigious institutions (Merton, 1973), while creating a vicious cycle for less prestigious institutions excluded from the allocation process of resources and status.

Therefore, the way higher education institutions are stratified based on the allocation of resources and privileges (especially from the government), as well as status in society, presents an enormous potential to affect some institutional characteristics and quality indicators commonly used in evaluations. For that reason, Astin (1985) criticizes quality efforts that conceptualize quality as a function of reputation and/or educational resources because they focus too much on available resources and all the other things resources can bring to a college, rather than looking for better evidence of educational quality.

The analysis of the effort to assess quality in higher education using the lens of institutional stratification opens up the possibility to explore how educational quality indicators are distributed within and between sectors of the Brazilian Higher Education system. Specifically, the inclusion of institutional characteristics makes it possible to investigate the degree to which student performance and other quality indicators assessed by ACO are related to the institutional type (university vs. non-university), the type of

administrative control (public vs. private), program selectivity, and socio-economic status (SES) of the student body.

Cultural Capital

The concept of cultural capital from Bourdieu's (1973) theory of social reproduction has been widely used to explore the influence of family background in education. Cultural capital in this theory is defined as the set of values, beliefs, and attitudes drawn from the dominant social class (Nogueira & Nogueira, 2002). These values and beliefs are assumed to be the norm for society; therefore, they are likely to be selected as the basis for developing tests such as the ENC. Because of that, students from dominant social classes are expected to benefit from the cultural capital they have already acquired, while students from other social classes have to struggle to acquire these values associated with college success.

Nogueira and Nogueira (2002) also discuss the relationship between socio-economic background and school performance. They point out that students from families with higher SES background are more likely to achieve higher performance in college. In the same line of thought, work by Thomas (1998) proposes to use student socio-economic background as a proxy for pre-college preparation and cultural capital.

The notion of cultural capital (Bourdieu, 1973) provides support for analyzing how student socio-economic background, represented in this investigation by parental income and educational level, is associated with student performance on the ENC. This influence is particularly critical when analyzing student performance in college. This is

because students come to college with very diverse individual characteristics and experiences, usually associated with the socioeconomic background that research shows affects academic achievement in college.

Methodology

Quantitative methods were used to address the research questions posited. The methods included descriptive analyses, correlation analyses, cross-tabulation analyses, one-way ANOVA, Kruskal Wallis test, and multiple regression technique. Initially, descriptive analyses of the variables at institution level, program level, and student level were conducted to present institutional characteristics, program characteristics, and student characteristics.

The first research question was examined by carrying out correlation analyses, descriptive analyses, and cross-tabulation analyses to explore the relationship between program performance (defined as mean student performance on the test), faculty characteristics, instructional resources (laboratories and libraries), curriculum characteristics, program selectivity, and institutional characteristics.

The second research question was explored by carrying out a multiple regression analysis for each academic field (Administration, Law, Civil Engineering, and Dentistry). This analysis was conducted to analyze how student performance on the ENC was associated with individual variables, program variables, and institutional variables. In addition, a series of blocked regressions was conducted to examine how the effect of institutional characteristics, program evaluation ratings, and student SES background changed as other variables were included in the regression equations.

Scope of the Study

Although the Brazilian educational evaluation system managed by the Ministry of Education includes three evaluation systems for elementary and secondary education, undergraduate education, and graduate education, the scope of this investigation is limited to the analysis of the evaluation system for undergraduate programs. Regarding the statistical analyses addressing the relationship between the outcomes of the ENC and ACO, the scope is limited to the following academic fields: Administration, Law, Civil Engineering, and Dentistry. In addition, only the programs from these fields that had been evaluated by ACO between 1997 and 1998 and whose students participated in the ENC in 2001 were included in the statistical analyses carried out in the present investigation. The decision to narrow the scope of this investigation was undertaken to make the data manageable while making it possible to analyze the relationship between the outcomes of ENC and ACO in different academic fields.

Limitations of the Study

This investigation presents some limitations. The most important ones are related to the scope of this investigation, the period of analysis, the data available, the generalization of the findings, and the structural differences between the Brazilian and American higher education systems.

The first limitation relates to the fact that this investigation examines this evaluation system at the national level. The choice of investigating this evaluation policy at the national level was taken because there is a lack of comprehensive studies exploring

the underlying assumptions of this evaluation system and its implications to the higher education system. Even though there is a great need for such a study, this investigation is limited due to the fact that it does not conduct any analysis at the institutional level, which would help reveal how faculty, administrators, and students view this evaluation system and how they have responded to its implementation. Nevertheless, this investigation certainly provides the theoretical basis for conducting further investigation at lower levels.

The second limitation relates to the period selected to explore the relationship between the outcome of the ENC and the outcomes of ACO. Although the results of the ENC are available for the period between 1996 and 2002, only one cohort (2001) was selected. The decision to concentrate efforts on only one cohort is justified in that it made it possible to carry out an in-depth analysis of four academic fields; this provided very interesting insights into how the relationship between the outcomes of this evaluation system was related in fields with different structural characteristics.

The third limitation is related to the data available to explore the relationship between the outcome of the ENC and the outcomes of ACO. While the national database maintained by the INEP (National Institute for Educational Studies and Research) offered an excellent opportunity to conduct this investigation, it also presents some limitations related to the nature and specificity of the data collected. The socioeconomic questionnaire (applied to students at the same time they take the ENC) was not designed for this investigation and, therefore, does not collect all data necessary to address important questions, such as those related to the good academic practices in higher

education. Furthermore, there is no information about student pre-college performance, which makes any inference about the effectiveness of individual programs impossible. Despite the fact that this questionnaire also collects some data about student SES background, this information reflects student SES condition at the end of their program. Therefore, it is necessary to take into consideration that the SES might have changed during the period students were attending college, although it is not expected that there were a lot of changes.

The fourth limitation is associated with the ability to generalize the findings. On the one hand, the findings from this investigation should not be generalized to the academic fields excluded from this investigation. Due to the size of the data and the complexity in performing the statistical analyses, the following academic fields were selected: Administration, Law, Civil Engineering, and Dentistry. These fields might significantly differ in some aspects, such as program selectivity, students' preparation and SES background, faculty qualification, infrastructure, and so forth. On the other hand, only those undergraduate programs with information available from the ENC and ACO were selected. Using this criterion, about two-thirds of all undergraduate programs in the selected academic fields were included in this investigation. Therefore, the findings of the quantitative part of this investigation may not be directly generalized to other undergraduate programs and academic fields not included in this investigation. Further investigations need to extend this investigation to other cohorts, academic fields, and undergraduate programs in order to assess whether the findings of this investigation can be extended to all higher education programs and other academic fields.

Additional limitations are related to the some structural differences between the Brazilian and American higher education systems. The reader must keep in mind at least the following peculiarities of Brazilian Higher Education that make it considerably different from the American higher education system. The assessment culture among Brazilian colleges and universities is in its initial stage, especially among small private colleges. This means that the vast majority of institutions of higher education do not have either the evaluation culture or the institutional capabilities to carry out the assessments as they are currently conducted in American colleges. Further differences relate to the participation of private institutions and the function and importance of public universities in Brazilian Higher Education. Federal and some state universities maintain a quasi-monopoly of graduate education and research, which gives them an overwhelming advantage over private institutions (especially small colleges) in terms of the quality of education and resources available. Despite the fact that private colleges and universities enroll over two-thirds of undergraduate students, most of them are small colleges that admit the least prepared college students. Therefore, the influence of public institutions in the formulation of this evaluation policy as well as the results of the evaluation instruments (ENC and ACO) might be exactly the opposite of what would be found in the United States. Careful examination of the findings and implications demands that the reader keep these differences in mind in order to derive the appropriate meaning from the results of this work.

Organization of the Study

This investigation is organized into four more chapters. Chapter II presents the literature review about program theory, quality and excellence in higher education, and monitoring quality in higher education. The review of the literature also includes works about monitoring college performance, institutional stratification, and cultural capital that guides the examination of the relationship between the outcomes of this evaluation system. Chapter III explains the research methods used in this investigation. Chapter IV presents the findings derived from the statistical analyses of the data collected and produced by the evaluation instruments (ENC and ACO). Chapter V analyzes the findings presented in chapter IV, using the frameworks presented in chapter II. One aspect that is addressed is the implications of using the data provided by the evaluation instruments (ENC and ACO) to make inferences about undergraduate program quality and to guide institutional actions. Furthermore, recommendations for further research are provided, along with recommendations for improving this evaluation system.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

The purpose of this chapter is to review relevant literature necessary to build conceptual and theoretical frames to explore the research questions posited. This chapter reviews the literature related to quality in higher education, models for monitoring outcomes in higher education, institutional stratification, and cultural capital. This literature was useful for this investigation because it offers interesting insights that enriched the analyses of the research questions by providing alternative frames to better understand the outcomes of the Brazilian evaluation system for undergraduate programs as well as the relationship between them.

Literature Regarding Quality in Higher Education

Policies addressing quality are important mechanisms to promote changes in higher education (El-Khawas, 2002). A number of authors have discussed the implementation of policies addressing quality in higher education institutions in many different countries around the world during the last decade (Brennan & Shah, 1997; El-Khawas, 2002; Harvey, 1998; Woodhouse, 1995). Harvey (1998) observes, "Central to this process [quality system] is the emphasis placed on quality as a vehicle for delivering policy requirements within available resources" (p. 238). He then points out that governments around the world see quality systems as tools to enhance the relevance of

higher education, increase access, and ensure comparability among higher education institutions. Bogue (1998) argues that quality systems have embedded important governing and ideals of quality, including evidence of performance, quality improvement, accountability, and clarifying institutional missions. Due to its enormous potential to reshape specific sectors of higher education or even the whole system, Woodhouse (1995) discusses important issues associated with the development of a quality system, such as the definition of quality, the purpose of the quality system, the advantages of certain quality approaches, and the appropriateness of a single approach to all sectors of higher education, among others.

The review of the literature addressing quality in higher education is divided into three parts. The first part focuses on the importance of the purpose of the educational policy addressing quality issues. The second part presents relevant works addressing the multifaceted, multidimensional notion of quality, as well as different quality conceptualizations applied in higher education. The third part reviews relevant studies presenting the most common approaches used to monitor and promote the purposes of the quality policy.

Purpose of the Quality Effort

Some researchers address the importance of appropriately defining the purpose of the quality policy in higher education (El-Khawas, 2002; Frazer, 1992). In order to have a clear definition of the purpose of the quality effort, policymakers should take into consideration some aspects that are likely to influence the policy goals, such as the

current stage of the evaluation or quality culture in the higher education system (Gaither, 1998), as well as the problems to be addressed, the priorities of the higher education system, and the institutional capacity necessary to carry out the activities required to achieve the policy goals (El-Khawas, 2002).

Some authors have pointed out that quality systems around the world have been implemented by governments to respond to a wide variety of pressures. Billing and Thomas (2000) and Rear (1994) discuss, for instance, that educational policies addressing quality in higher education have emerged to deal with increasing enrollments and participation rates, increasing student diversity, changes in information technology, movement toward increasing institutional autonomy, pressure for public accountability, pressure for rationalizing spending, overseas provision of education, and curricular flexibility, among others.

Responding to the pressures for acting to promote certain changes in higher education, governments have developed educational policies to achieve a wide range of purposes. Most of the quality policies have been implemented to enhance educational quality, maintain educational quality, and provide accountability (Burke & Minassians, 2002; Frazer, 1992; Harvey, 1998; Lim, 2001; Patrick & Stanley, 1998; Vroeijenstinjn, 1992; Woodhouse, 1995). El-Khawas (2002) states that quality assurance systems have also been employed, aimed at "raising standards or ensuring quality control; correcting deficiencies or avoiding unacceptable behavior" (pp. 205-206). Furthermore, some research mentions educational policies developed to gather information for planning and resource allocation (Frazer, 1992, Vroeijenstinjn, 1992; Woodhouse, 1995), enhance

flexibility of higher education institutions in regard to new forms of curriculum, teaching, and learning strategies (Lim, 2001), and increase control over higher education institutions (Bogue, 1998; Harvey, 1998).

Most of the time, quality policies have been developed to address multiple purposes. Bogue (1998) and Harvey (1998) point out that the development of policies in some countries was associated with the need for expanding access and providing public accountability. Lim (2001) and Lively (1992) make the connection between accountability and the increasing demand for public appropriations in a period of financial constraints. Frazer (1992), Gaither (1998), Harvey (1998), Neave and Van Vught (1991), and Vroeijenstinjn (1992) point out that quality systems have demanded more accountability in exchange for increasing institutional autonomy. Gates and associates (2002) state that accountability systems have been used to provide information for stakeholders and assist program funding, planning, improvement, and elimination.

The literature reviewed is not conclusive with regard to which of these purposes would better serve the cause of higher education. Harvey (1998) argues that the primary purpose of a quality policy should be to improve the overall quality of the higher education system. Others argue that it is possible to have a quality system designed to promote both improvement and accountability (Burke & Minassians, 2002; El-Khawas, 2002). In contrast, Gaither (1998) raises an important issue related to whether the main purpose of a quality assurance system should be accountability or program improvement. Harvey (1998) discusses the tension between accountability and improvement, while

Ewell (1994) doubts the compatibility of promoting improvement and accountability at the same time.

Quality Conceptualization

A great deal of higher education literature addresses quality conceptualization as an important issue in the policy formulation process (Astin, 1985; Bogue, 1998; Harvey & Green, 1993; Nordvall & Baxton, 1996). The way quality is conceptualized provides the philosophical basis for choosing the approach used to address quality (Barnett, 1994) and the outcome measures (Barnett, 1994). As Astin (1991) points out, "The concepts of excellence are usually incorporated in policies and practices, representing the values that guide the institutional effort to promote excellence" (p. 5). For that reason, it is important to examine different perspectives commonly used to conceptualize quality in higher education, focusing on the multifaceted notion of quality, the stakeholders' perspectives of quality, and the multiple approaches used to conceptualize quality.

A number of scholars have addressed the multifaceted notion of quality in higher education (Bogue & Saunders, 1992; Conrad & Pratt, 1985; Frazer, 1992; Harvey & Green, 1993; Van Vught, 1994; Vroeijenstinjn, 1992). Van Vught argues for using various interlocking, cross-cutting criteria when assessing quality in higher education. Harvey and Green address the range of qualities in higher education, while Vroeijenstinjn argues that it would be more appropriate to characterize quality in higher education as having multiple characteristics, suggesting qualities (plural) rather than quality (singular)

of higher education, due to its complex and multifaceted nature. Bogue and Saunders summarize the claims of the multifaceted notion of quality by observing,

> Quality is not a single-factor attribute of personal or organizational performance.... [N]o single indicator of measurement of quality can stand without some criticism.... because individual and institutional performance are, as we have noted, too complex to be captured in a single point of evidence. (pp. 220-221)

Due to the multifaceted notion of quality in higher education, some studies address the difficulty in developing a unique, accurate, unambiguous notion of quality in higher education. Higher education stakeholders might have completely different concepts of quality depending on their values and interests in higher education (Chen, 1990; Harvey & Green, 1993; Lim, 2001; Vroeijenstijn, 1992). Even more complex is the fact that sometimes the same stakeholder conceptualizes quality in a different way depending on the situation (Harvey & Green, 1993). Harvey and Green go further to claim that "this is not a different perspective on the same thing but different perspectives on different things with the same label" (p. 10). This observation calls attention to understanding different approaches stakeholders might use to address quality in the way they perceive its importance for the higher education system.

A number of scholars have presented different approaches to characterize quality in higher education. Harvey and Green (1993) state that quality in higher education can be characterized as exception, as perfection or consistency, as fitness for purpose, as value for money, and as transformative. Astin (1985), Bogue (1998), and Nordvall and Baxton (1996) describe other approaches to define excellence and quality in higher

education. They are the reputational approach, the resources view, the outcomes view, the value-added approach, and the fitness for purpose approach.

Despite the multiple definitions of quality, Harvey and Green (1993) point out that these approaches are intercorrelated. For instance, Bogue (1998) observes that the reputational approach and resources approach are based on resources, prestige, selectivity, institutional differentiation, and program diversity. Burke and Minassians (2002) refer to a joint approach named resource-and–reputation approach, while Harvey and Green (1993) offer the approach to define quality as exceptional, which uses the same values as the reputational and resources approaches.

Due to the complexity and multifaceted characteristic of quality, several works in the education literature have criticized some of these approaches previously presented. The approaches that have been more frequently criticized are reputation, resources, and quality as exception, which conceptualizes quality as a function of resources, prestige, selectivity, and institutional differentiation. Lim (2001) points out that these approaches embody an elitist view of higher education, where only a few institutions would be able to attain high quality, which requires "huge excess demand and limited places" (p. 14). By the same token, Bogue (1998) offers the theory of limited supply to explain that quality assumes a pyramid structure under which excellence can be achieved by selective institutions. Burke and Minassians (2002) criticize the reputational and resources approach for focusing on inputs (funding, students, faculty) and disregarding the quality of services provided for students, states, and society. Furthermore, Astin (1985) states, "Just because people believe that the institution is excellent, it does not necessary follow

that the institution is any more effective in developing talent than a less prestigious institution" (p. 50). Frazer (1992) also rejects this approach with the defense that "quality cannot be equated with excellence" (p. 15).

Based on the critique of the traditional approaches, Astin (1985) and Nordvall and Baxton (1996) offer two alternative approaches to overcome some problems: the talent development approach and the course-level approach. In general, the talent development approach (Astin, 1985, 1991) is similar to the value-added approach; however, the main focus is on the talent developed by students during college. The course-level approach (Nordvall & Baxton, 1996) for defining quality of undergraduate education focuses on course-level academic processes. The focus of attention is on daily activities carried out by faculty and students that are more directly related to the teaching-learning process.

Quality Approaches

A vast amount of literature has presented some approaches used in educational policies addressing quality around the world. The major distinction among the quality approaches relates to the goals of the quality policy, the methods used to promote the policy goals, and the control or coordination of the process (Stanley & Patrick, 1998; Westerheiden, Brennan, & Maassen, 1994).

The most common approaches employed to address quality in higher education are quality control, quality assurance, quality audit, and accreditation, as presented by Frazer (1992), Lim (2001), Gaither (1998), Patrick and Stanley (1998), Tam (2001), and Woodhouse (1995). Quality control is usually carried out by an external examiner system

to "verify whether teaching and assessment are carried out in a satisfactory manner" (Patrick & Stanley, 1998) or to "check whether the products produced or services provided have reached the pre-defined standards" (Tam, 2001). Quality assurance involves policies and processes implemented to ensure the maintenance and enhancement of educational quality (Gaither, 1998; Patrick & Stanley, 1998). Quality audit is defined by Frazer (1992) as the process that verifies whether the controls established by the quality control and quality assurance processes are working properly. Accreditation is usually carried out by an external body to ensure that institutions have defined appropriate educational goals and have implemented the necessary conditions to achieve the goals, and whether the goals have been accomplished (Chernay, 1990, as cited in Frazer, 1992).

Other approaches have also been analyzed in the literature. Some literature describes quality assessment (Gaither, 1998; Patrick & Stanley, 1998), peer review (Frazer, 1992; Van Vught, 1994), program validation (Woodhouse, 1995), and performance indicators (Harvey & Green, 1993; Tam, 2001). Harvey (1998) points out that joint approaches to quality in higher education have also used a combination of self-assessment, peer review, and statistical or performance indicators.

A number of studies suggest integrating internal procedures to maintain and enhance quality, with external procedures to monitor quality (Burke & Minassians, 2002; Harvey, 1998; Lim, 2001; Van Vught, 1994; Vroeijenstijn, 1994; Woodhouse, 1995). Gaither (1998) observes that the implementation of educational policies that exchange institutional autonomy for accountability requires "a delicate balance between internal

improvement and external accountability" (p. 2). Burke and Minassians (2002) defend the need for having accountability tied to quality improvement, where the main focus would be on performance instead of compliance. Lim (2001) suggests that internal and external procedures should be used to audit the quality of processes and the achievements. In fact, some of the approaches previously presented have embedded a mix of internal and external processes to deal with quality, such as quality assurance and quality audit (Gaither, 1998; Lim, 2001; Patrick & Stanley, 1998). Van Vught (1994) justifies his preference for a joint approach that includes self-evaluation coupled with peer review, arguing,

> A quality assessment system that only relies on collegial peer review without any reference to the needs outside of the higher education system, implies the risk of an extreme isolationism of the higher education institutions from the rest of society (and thus the danger of the denial of legitimacy of their existence). A quality assessment system which is limited to only providing accountability to external authorities denies some of the basic organizational characteristics of higher education institutions and therefore implies the risk of not being taken seriously by the professional experts. (p. 39)

Despite the claim for developing quality systems based on a combination of self-regulated and externally regulated approaches, some authors have stressed the importance of having quality systems whose main purpose is quality improvement. Gaither (1998) observes a movement from quality management toward quality enhancement through moving beyond mere measurement and maintenance. Harvey (1998) shares a similar view about the future of quality policies in higher education, urging for a paradigm shift from accountability toward implementing a culture propitious for the improvement of process. He argues that "higher education needs to produce transformative agents—critical reflective citizens—and the external quality monitoring must help, not hinder, that

development" (p. 14). Harvey (1998) defends his point of view by arguing that external quality monitoring has only initial impact on quality improvement, while self-evaluations have been shown to be more effective in promoting quality improvement. He goes further, arguing that even in countries that have successfully implemented external quality monitoring, there are some doubts about the development of a culture necessary to promote continuous improvement.

Some literature urges implementing a self-regulatory approach (Harvey, 1998) or a self-evaluation (Westerheiden, Brennan, & Maassen, 1994) to deal with quality in higher education. Harvey points out that the self-regulatory approach would deal better with some concerns related to the high costs of external monitoring procedures, as well as to the threat to academic freedom and autonomy of universities presented in other approaches. He also mentions that self-assessments followed by peer review provide a better opportunity for exploring institutional purpose, effectiveness, weakness, and opportunities, which would foster the development of an institutional culture necessary for promoting continuous improvement. Westerheiden, Brennan, and Maassen (1994) defend self-evaluation as the method for quality assessment, where "internal aspect is externally supplemented with a *site visit* by a team of experts, often an academic peer review committee" (p. 19).

Some works have addressed the effectiveness of quality approaches in higher education. Woodhouse (1995) calls for the development of institutional quality systems that are efficient and effective by ensuring appropriate quality control and assurance procedures, checking whether the standards are feasible, and reducing the number of

procedures involved in all levels of the quality system. Sharing similar ideas, Harvey (1998) observes,

> The effectiveness of external monitoring depends on three things: the withering away of the bureaucratic, accountability, conformance process; the linking of a lighter-touch external review to well-developed internal procedures for quality improvement; the development of an internal quality culture, widely embraced, for which internal procedures are guides and aids to appropriate practice. (p. 10)

On the other hand, Gaither (1998) stresses the importance of faculty involvement in the establishment of the quality system, arguing that "quality is best maintained and enhanced through the professional commitment of the faculty, not through quality assurance systems, controls, or legislation" (p. 1).

Another concern present in the literature refers to the implications of having a joint quality approach involving multiple agencies. Woodhouse (1995) discusses the advantages associated with a single agency and with multiple agencies. Despite some advantages associated with a quality system whose control is shared among multiple agencies, Woodhouse suggests developing a quality system coordinated by a single agency to "ensure complete coverage of the desired quality features; avoid unintended overlaps; have deliberate overlaps only for triangulation or cross-checking; minimize the load on the institutions; and maximize the efficiency of the external system" (p. 7). He goes further to imply that joint institutional ownership of the agency (inter-institutional system for self-regulation) would increase the likelihood of enhancing quality because institutions of higher education by themselves are responsive to institutional quality and would reduce the costs associated with external quality monitoring.

Lastly, due to some peculiarities of some higher education systems, some literature is concerned with the process of borrowing quality systems from one country to another country. Rhoades and Sporn (2002) analyze the development of quality policies in the United States and European countries, focusing on the influence of global professional and political economies. El-Khawas (2002) raised concerns with the problem of borrowing quality systems adopted in other countries due to "distinctive circumstances and infrastructure, and with differing educational traditions" (p. 199). Billing and Thomas (2000) discuss some problems related to international transferability of quality systems, focusing on cultural, structural, political and technical issues.

In summary, the literature on quality in higher education addresses important issues involved in the process of designing and implementing an educational policy to deal with quality in such complex types of organizations as higher education institutions. From the policymaking standpoint, the definition of an adequate purpose of the quality effort is relevant because it affects the design, implementation, and outcomes of the quality system. Due to the multifaceted notion of quality in higher education, the literature presents many approaches used to define quality; each one has embedded its own values and ways to promote quality that produces different educational outcomes and benefits individuals and institutions differently. Based on the complexity of the quality in higher education and the multiplicity of conceptualizations, the literature addresses the importance of designing a quality system that creates a propitious environment for enhancing the quality of educational processes and outcomes and, at the same time, creating a mechanism for keeping higher education institutions accountable.

Models for Monitoring Outcomes in Higher Education

The literature on quality in higher education indicates an increasing interest in monitoring and assessing educational quality as a result of the recent movement to implement quality assurance policies in higher education. Even in those quality systems where accountability is the main purpose, there is an emerging movement that calls for a shift from regulation to results (Burke & Minassians, 2002). In order to provide accurate information about the effectiveness of an educational policy addressing quality of undergraduate programs, it is necessary to have a frame for analyzing the educational outcomes.

Some literature provides basic guidelines for monitoring and assessing quality in higher education and educational outcomes. El-Khawas (2002) argues that quality assessment in higher education requires a more comprehensive approach that involves the analysis of sufficient capacity, efficiency, and effectiveness of higher education institutions. Other scholars claim that any process designed to assess quality of teaching in higher education should include a comprehensive approach involving inputs, processes, and outputs (Astin, 1991; Burke & Minassians, 2002; Frazer, 1992; Harvey & Green, 1993; Vroeijenstinjn, 1992). Astin (1991) goes further to refer to outcome-only assessments, environment-outcome assessments, input-outcome assessments, and environment-only assessments as incomplete designs for assessing activities in higher education.

Willms (1992) proposes the Input-Process-Output Model (IPO) for monitoring school performance based on a multilevel approach that includes inputs, processes, and outputs. He presents the rationale for this model, stating that

> monitoring systems are based on a theory that schooling 'outputs', such as academic achievement, attitudes, and aspirations, are predominantly determined by pupil 'inputs', including sex, race, ability upon entering school, and family background, and by the many school-related 'processes' that determine a school's context and its inner workings. (pp. 144-145)

Despite the fact that this model includes controls for student background characteristics, it is limited in the sense that these controls lack the basis for considering the influence of some school characteristics (labeled as contextual factors) over inputs, processes, and outcomes (Bosker & Scheerens, 1994). For instance, Bosker and Scheerens are concerned about the role of school selectivity in attracting good students, good teachers and good administrators. Based on this concern, other scholars advocate for a shift from input-process-output approach to a more comprehensive approach called contextual-input-process-output approach (Creemers and Scheerens, 1994; Bosker and Scheerens, 1994).

The need for developing a theoretical framework to examine effectiveness and quality in higher education has also been a concern among scholars in higher education. Astin (1970, 1991, 1993) developed the Input-Environment-Output model (IEO) for assessment of activities in higher education based on the talent development approach previously discussed. He defines the basic elements of this model:

> Inputs refer to the characteristics of the student at the time of initial entry to the institution; environment refers to the various programs, policies, faculty, peers, and educational experiences to which the student is exposed; and outcomes refer to the student's characteristics after exposure to the environment. (Astin, 1993, pp. 144-145)

In fact, the IEO model was designed to assess the effects of educational environment on outputs by controlling for input differences, which reduces potential biases associated with preexisting differences among students (Astin, 1991). However, the concern expressed by Bosker and Scheerens (1994) with regard to the inclusion of contextual aspects in the IPO model also applies to the IEO model.

A number of studies have applied Astin's IEO model to explore the influence of input characteristics and environment variables on diverse educational outcomes. Hu and Kuh (2003) explore a learning productivity model, which is based on Astin's IEO model and Pascarella's (1985) causal model, to examine how institutions affect student learning. House (1999) examines the influence of entering characteristics and college experiences on student satisfaction and degree completion. Young and Fischer (1996) investigate the influence of undergraduate and post-college activities and experiences on alumni giving. Others explore the relationship between environmental variables and college environment and changes in students' religious beliefs and convictions (Lee, 2002), enrollment in science, math, and engineering graduate programs (Sax, 2001), graduate school attendance (Walpole, 2003), and student satisfaction in Web-based courses (Thurmond, Wambach, & Connors, 2002).

To sum up, an increasing number of governments have implemented educational policies to monitor higher education institutions' responses to strong pressure to provide accountability with regard to the quality of educational outcomes. The models presented in this section clearly stress the importance of considering multiple dimensions involved in and associated with educational outcomes. A great deal of research argues for

considering the characteristics of educational inputs and educational processes as fundamental components for the analysis of the outputs. In addition, the inclusion of contextual factors in higher education makes it possible to investigate the degree to which certain characteristics of higher education institutions (e.g., institutional type, administrative control, and institutional selectivity) affect student performance in college, as defined by Hu and Kuh (2003) as environment variables. Additional attention is required not only to select the model to explore the educational outcomes but also to recognize the characteristics and limitations of the educational outcomes. This calls attention to the use of standardized tests in higher education as a measure of educational output, which is provided in the following section.

Use of Standardized Tests in Higher Education

A number of researchers have discussed the use of standardized tests to assess student outcomes in American higher education. Pascarella and Terenzini (1991) review a large number of studies that have employed some form of standardized tests to measure learning and development during college. Astin (1993) points out that performance on standardized tests has been employed in the U.S. for college admission and college placement (e.g., SAT, ACT), professional certification (e.g., NTE), and for admission to graduate or professional school (e.g., GRE, MCAT, LSAT), among other things. Ewell (2001) discusses the increasing trend among states in the U.S. to employ standardized tests in higher education institutions to (1) certify student readiness or achievement (e.g., Texas and Florida), (2) induce particular kinds of institutional behaviors aiming to

improve quality (e.g., Tennessee and Missouri), and (3) demonstrate accountability (Utah, South Dakota, and Arkansas).

The increasing use of standardized tests to assess educational outcomes calls attention to the discussion of the type of outcomes these tests are designed to measure. A number of approaches have been proposed to classify and assess student outcomes in higher education (Astin, 1974, 1977, 1993; Brown & DeCoster, 1982; Lenning, Lee, Micek, & Service, 1977; Jacobi, Astin, & Ayala, 1987). Astin's taxonomy has been widely used to classify student outcomes in higher education (e.g., Dey & Associates, 1997; Pascarella & Terenzini, 1991). It classifies outcomes into two domains based on the type of data (psychological and behavioral) and the type of outcome (cognitive and affective). The literature on student development and assessment in higher education refers to academic ability, critical thinking, basic learning skills, special aptitudes, and academic achievement as psychological cognitive outcomes and values, interests, self-concept, attitudes, beliefs, satisfaction with college, involvement, and effort as psychological affective and motivational outcomes (Alexander & Stark, 1986; Astin, 1993, 1999; Bogue & Saunders, 1992; Dey & Associates, 1997; Pascarella & Terenzini, 1991). This taxonomy is relevant to the application of standardized tests in higher education because these tests have been employed to assess internal traits of students that cannot be directly assessed otherwise (e.g., psychological outcomes).

A number of studies (Bastos, Michael, & Oliveira, 2001; Martuza, 1977; Mehrens & Lehmann, 1987; Popham, 1999; Schmoker, 2000; Stufflebeam, 2001) discuss the types of standardized tests have been used to assess psychological outcomes in education:

norm-referenced tests (NRTs) and criterion-referenced tests (CRTs). Regardless of the type of test, Mehrens and Lehmann (1987) urge that standardized tests should provide information to help make decisions in four domains (e.g., instructional, guidance, administrative, or research) that affect individuals (students and teachers), classrooms, schools/colleges, or whole educational systems. Since the Brazilian national evaluation system for undergraduate programs utilizes an NRT as its main component (ENC), this type of test deserves more attention in the remaining review of the literature.

Martuza (1977) points out that NRTs are recommended to analyze factors that are normative in fashion, such as achievement, aptitude, and some specific personality factors. Mehrens and Lehmann (1987) claim that NRTs are suitable for providing a broad view of achievement levels, for identifying strengths and weaknesses of students and schools by comparing them to others (see also Popham, 1999), for providing evidence about the effectiveness of a given program in comparison to others, for promoting maximum development and recognizing outstanding performance, and for selection decisions in special situations where there is a fixed-quota selection decision process.

A number of researchers have discussed the limitations of NRTs in assessing educational outcomes. NRTs have been criticized for (a) being too narrow to judge the program's worth and merit (Stufflebeam, 2001); (b) assessing only very specific skills that overlap very little with other outcome measures (Astin, 1993); (c) not providing information about what and how much the student has learned (Mehrens & Lehmann, 1987); (d) not providing information for program improvement (Martuza, 1977; Popham, 1999; Stufflebeam, 2001); and (e) ordering students by aptitude rather than by knowledge

developed in school (Bishop, Manã, & Bishop, 2001). Stufflebeam (2001) also presents other limitations associated with this type of test, stating that

> they provide data only about student outcomes; they reinforce student's multiple-choice test-taking behavior rather than their writing and speaking behaviors; they tend to address only lower-order learning objectives; and they are perhaps a better indicator of the socioeconomic levels of the students in a given program, school, or school district than of the quality of teaching and learning. (pp. 21-22)

Other scholars have also discussed potential biases associated with the interpretation of the results of the NRTs. Martuza (1977) provides test designers and policymakers with a comprehensive analysis of the reliability and validity of NRTs (and CRTs as well), which includes some situations that may result in biased outcomes of these tests and, therefore, may threaten the meaning and usefulness of the tests. Bogue and Sounders (1992) and Flaugher (1974, as cited by Bogue & Sounders, 1992) discuss potential biases incorporated in standardized tests associated with previous experience, language, culture, and gender as well as the biases in test content, test Administration, and test use. Kohn (2000) and Shepard (1981) concern about the fact that the differences in test-score performance may be associated with racial and socioeconomic background. Furthermore, Popham (1999) acknowledges that differences in performance in NRTs are due to what is taught in school, native intellectual intelligence, and out-of-school environment associated with stimulus-rich environments and socioeconomic factors so that the test score cannot reflect exclusively school effectiveness. In addition, Astin (1999) mentions that the use of norm-based testing has raised equity issues related to fair competition for disadvantaged groups of students and the negative feedback to students who receive low test scores about their performance and capabilities.

Astin (1991, 1993) discusses contradictory findings about the use of standardized tests to assess and promote student outcomes. He found that standardized tests relate to certain environmental factors and involvement experiences that do not affect other important student outcomes. This finding has significant implications for policymaking because it provides guidance to develop institutional actions that would result in higher test scores but not in better student development. Astin (1993), then, argues,

> The point here is not necessarily to denigrate the use of standardized multiple-choice tests in assessing educational outcomes—such tests are clearly of value in assessing certain kinds of student performance—but rather to point out that they measure rather narrowly defined skills and do not appear to be good indicators of student development in many important areas. (p. 429)

Another important issue relates to coaching students for taking standardized tests aimed at improving their test scores (Slack & Porter, 1980; Jackson, 1980; Messick, 1981; Popham, 2001). Such efforts might be associated with improving abilities and skills measured by tests, enhancing test-taking sophistication or reducing anxiety often associated with taking tests, and teaching answer-selection tricks. While the first two strategies do not threaten the validity of the test, the latter might result in a higher score not associated with abilities and skills measured by tests. As Koretz and Barrom (1998), Messick (1981), and Popham (2001) point out, the latter strategy might produce fLawed results due to unacceptable behaviors, such as teaching for the test and cheating.

In addition to the possibility of artificially inflating the test scores, there is another controversial issue associated with applying for college. McDonough (1994) analyzes how upper–middle-class students are getting involved in an "industry" to help them to apply for college that includes coaching students to take the SAT test. Bond (1981) then

expresses his concern about the consequences of such processes on students from lower-SES background, observing that "if commercial coaching schools are effective in raising test scores, then economically advantaged students, who can afford such schools, have yet another advantage over their less fortunately situated cohorts" (p.65).

The literature review about the use of standardized tests reveals important concerns that policymakers and educators should be aware of. The results of NRTs should be used carefully and in accordance with the test objective, which is primarily to order individual students (or programs) according to the performance of all students who participated in the test. Nevertheless, a great deal of literature addresses increasing concern about limitations, potential biases, and unfairness of NRTs, as well as unacceptable behaviors, to artificially raise the test scores. This literature provides a useful framework not only to guide the analysis of the results of this investigation but also to provide information about the problems and limitations associated with NRTs that might threaten the interpretation of the results.

Institutional Stratification

Trow (1984), in his work "The Analysis of Status", provides the perspective of stratification to explore the mechanisms by which institutions of higher education are stratified in terms of status and prestige, wealth, power, or some combination of these aspects. Under this perspective, there are two principles of hierarchy through which higher education institutions are stratified within and between sectors: "the competitive advantage in the market for the factors that make for academic prestige" and "the

allocation of functions, rights, privileges, and resources by governments to specific institutions or sectors" (p. 162).

Trow (1984) defends the application of this perspective to understand the implications of the stratification of the higher education system. He states,

> The perspective of stratification in higher education brings us very quickly to more substantial issues, such as the organization of higher education systems, the relation of status to function, the academic division of labor, institutional autonomy, faculty power and authority, and the quality of teaching and research within institutions of various kinds and ranks. (p. 133)

Trow discusses the application of the perspective of stratification in many aspects of the higher education system; however, the review of the literature on institutional stratification focuses on the influence of the organization of the higher education system and the character of its stratification on some aspects related to the teaching process, as well as some educational outcomes. Specifically, this review addresses how some educational outcomes are likely to be influenced by institutional status, institutional selectivity, institutional type and control, joint-production of research and teaching, and research orientation.

Merton (1973, as cited in Trow, 1984) addresses an important mechanism by which institutional status and resources might exert an overwhelming influence on the educational outcomes in elite institutions. He explains how status, prestige, and resources of certain sectors of higher education are actually used as a source of increasing advantage for institutions that are already gifted. As has been observed in the United States, elite institutions tend to attract the best resources available for higher education (including students and faculty) because of their good reputation, financial resources,

qualified faculty, highly selective admissions process, and good infrastructure. This situation creates a virtuous cycle for the most prestigious institutions while creating a vicious cycle for less prestigious institutions, which are excluded from the process of allocation of resources and status. The Mathew effect, as defined by Merton, presents an enormous potential to affect educational outcomes, since the most academically prepared students tend to select the college to attend by taking into consideration the institutional prestige, selectivity, and quality of faculty and infrastructure.

The interest in investigating the effect of institutional type and institutional control on college outcomes is present in higher education literature. For instance, Astin (1993), Dey and associates (1997), and Stoecker, Pascarella, and Wolfe (1988) examine different college environments in terms of institutional type, student body selectivity, size, graduate emphasis, and financial resources. Pascarella and Terenzini (1991) review a large amount of work that investigates the effect that different kinds of postsecondary institutions have on student change and development during college. They found that the effect associated with organization and structural characteristics on a wide range of student outcomes was smaller than expected, having much more impact on socioeconomic-oriented outcomes than on developmentally oriented outcomes. They suggest that this finding challenges the notion of institutional quality based on resources, simpleminded outcomes, and reputation, arguing that "the quality of undergraduate education may be much more a function of what colleges do programmatically than it is of the human, financial, and educational resources at their disposal (Pascarella and Terenzini, 1991, p. 637)

Some authors examine the influence of institutional selectivity on educational attainment. Pascarella and Terenzini (1991) cite some work (e.g., Anderson, 1984; Clewell & Ficklen, 1986; Cope & Hannah, 1975; Fetters 1977; McClelland,1990; Perricci, 1980; Stoecker, Pascarella, & Wolfe, 1988; Tinto, 1987) that demonstrates the influence of institutional selectivity on student persistence, bachelor's degree completion, or the percentage of graduates moving on to graduate and professional school. Pascarella and Terenzini (1991) also explain how institutional selectivity brings an initial advantage to a college:

> In attempting to understand the influence of college "quality" on educational attainment, it is important to note that "high quality" colleges start with a distinctive advantage in terms of the academic ability, educational aspirations, level and clarity of career ambition, and family financial resources of the students they recruit and enroll (for example, Astin, 1982; Clark, Heist, McConnell, Trow, & Yong, 1972; Hearn, 1984; Karabel & Astin, 1975). (p. 374)

Other works examine the influence of institutional selectivity on student performance on standardized tests. Ewell (2001) and Astin (1999) are highly skeptical about the use of these tests to assess student outcomes because their results are highly affected by initial differences among students from institutions with different selectivity levels. Astin (1999) observes,

> In American higher education we have developed a set of elite institutions that are so selective in their admissions that high performance standards at exit are almost guaranteed, even if the institution contributes little to the educational process. Ironically, these same institutions have the best facilities and the most resources of all institutions (Astin, 1985). (p. 162)

Ewell (2001) agrees with Astin (1985)'s point of view about the influence of institutional selectivity, making an argument similar with that provided by Merton (1973) about the virtuous cycle for selective institutions.

> In cases where raw achievement scores are used to make consequential judgments about institutional performance, for instance, the most selective colleges and universities will inevitably benefit: the "rich will get richer" and the "poor" will be motivated to become more selective. (Ewell, 2001, p. 5)

Lastly, some works have addressed the impact of institutional research orientation on certain educational outcomes. Kim, Rhoades, and Woodard (2003) review the literature addressing the trade-off between teaching and research with regard to the emphasis on teaching, instructional activity, educational attainment, and student persistence. They found no evidence, either positive or negative, in the literature (e.g., Braxton, 1996; Feldman, 1987) to support the thesis for the existence of a trade-off between research and teaching effectiveness. In addition, they cite other works (Milem et al., 2000; Rhoades, 2000) that suggest the magnitude of this trade-off is smaller than is usually represented. In contrast, they refer to some works that suggest that instructional activities are being subsidized by research (Leslie, Rhoades, & Oaxaca, 1999) and efficiency is increased with the joint production of teaching and research (Colbeck, 1998; Dundar & Lewis, 1995).

Astin (1993) examined the influence of faculty emphasis on research on some educational outcomes. He found that faculty research orientation represented a positive effect on performance on standardized tests (e.g., LSAT), student's perception of institutional resources, and student's perception of reputation emphasis, but represented a negative effect on student orientation of faculty, satisfaction with faculty, perception of social change, leadership, growth in interpersonal skills, overall satisfaction with quality of instruction, and overall college experience, among others. He then suggests that institutions with strong faculty emphasis on research seem to pay a price for promoting

research to the detriment of promoting student development. A probable explanation for this is that faculty devote much more effort and time for research activities to the detriment of activities associated with teaching.

To summarize, the literature reviewed about institutional stratification offered the strongest framework within which to examine the outcomes of the evaluation system and, therefore, within which to make appropriate meaning of the results. The mechanisms by which higher education institutions are stratified presents a strong effect not only in the resources (both human and financial) they are able to attract, but also in the educational outcomes promoted. The "special help" some institutions are able to receive leads us to question how institutional effectiveness could be separated from such an influence. This framework then is very useful in addressing the need to analyze the results of this investigation in the appropriate context, trying to identify the mechanisms by which some (a few) institutions might benefit from the current evaluation system.

Cultural Capital

A great deal of literature addressing the influence of college on educational outcomes acknowledges the important role played by student socioeconomic background (SES) on several aspects of college life (Astin, 1993; Hearn, 1991; Pascarella & Terenzini, 1991; Tinto, 1985). As already discussed, institutional selectivity is associated with admitting students with better preparation for college, better educational outcomes, and higher student performance (Astin, 1985; Ewell, 2001; Pascarella & Terenzini, 1991). The other side of institutional selectivity is that it also brings more students from higher SES background; consequently, students from lower SES do not have the same

quality of opportunities and educational attainment as their counterparts from higher SES backgrounds.

Due to the influence of student SES background on almost all aspects of college life, including environmental measures, educational outcomes, and educational attainment, it is important to have a framework for exploring the way SES background works on college and impacts student outcomes. Although much of the research in higher education on student outcomes includes student SES background to control for a rival explanation, as summarized by Pascarella and Terenzini (1991), selecting a framework to explore the relationship between SES and student outcomes enriches the analysis.

The notion of cultural capital, drawn from Bourdieu's (1973) theory of social reproduction, provides a useful framework for exploring differences in educational outcomes among students from different social classes. Some research has used Bourdieu's theory as the framework within which to investigate the effect of cultural capital on educational attainment (De Graff, De Graff, & Kraaykamp, 2000), college experiences and outcomes (Walpole, 2003), participation in sports and academic achievement (Eitle & Eitle, 2002), and school grades (Dumais, 2002; DiMaggio, 1982). In addition, Persell, Catsambis, and Cookson (1992) examine whether the conversion of economic, social, and cultural-capital resources (usually conceptualized in terms of SES, academic ability, aspirations, and grades) are converted into higher educational attainment in the same way by men and women.

One important contribution of this theory relates to the analysis of social inequalities promoted by the educational system. Values, beliefs, and attitudes drawn

from the dominant social class are arbitrarily attributed as the norm for society (Freeman, 1997; Nogueira & Nogueira, 2002). By promoting the values of the dominant class, the educational system, then, acts to reproduce and legitimate the dominant values of society, which causes the perpetuation of social inequalities (Nogueira & Nogueira, 2002). Zweigenhaft (1994) presents the process under which the perpetuation of inequality occurs, stating that "those in power do not only merely pass their material wealth, or economic capital, to their offspring (though they do this), they also try to assure that their children acquire what he [Bourdieu] calls 'cultural capital' and 'social capital'" (p. 211). In addition, McDonough (1994) discusses the relevance of cultural captial for families with high cultural capital in their attempt to perpetuate the advantages through the educational system, stating,

> Cultural capital is an important form of capital and is often translated into elite educational credentials. Individuals from high cultural capital have clear investment strategies of how much and what kind of schooling they or their children should have. Parents with high cultural capital attempt to secure for their children as prestigious a college as possible because they know it will pay off in later job success and social status. (p. 430)

The adoption of values that only a select group of students have internalized creates a competitive advantage for some students, while creating a problem of cultural adaptation for other students. Dumais (2002) describes how the degree to which students fit the dominant culture affects higher-SES and lower-SES students differently:

> Children who have more cultural capital (having been exposed to it from birth in their upper middle-class families) feel more comfortable in school, communicate easily with teachers, and are therefore more likely to do well in school (De Graff, De Graff, and Kraaykamp, 2000). Lower-class students, on the other hand, find the school environment different from their home environment and lack the capital necessary to fit in as well as the higher-SES students. (p. 46)

Since students from low SES background are not exposed to the same social and familiar environments where dominant values are disseminated, they struggle to assimilate these values necessary to succeed in education (Dumais, 2002). As a result of cultural differences, students who lack the values rewarded in the educational system are likely to refrain from attending college; even when they go to college, they say that these students will not achieve the expected results (De Graff, De Graff, and Kraaykamp, 2000).

As seen in the brief review of the literature presented, cultural capital can be a useful tool to explore the influence of SES background on college outcomes. It is shown in the literature that students from low SES background differ from high SES peers by having different educational aspirations, getting less involved in college, reporting lower college grades, graduating in a lower proportion than students from high SES background, and benefiting less from college degrees. By framing the analysis of the effect associated with different SES background on student performance, it is possible to explore the degree to which students from different social classes have the same opportunity to attend college in different fields, as well as whether they are able to present an equivalent level of performance in the exit exam.

Summary

The literature review was carried out with two main intentions. The first was to review some work addressing conceptual characteristics necessary to develop a better understanding of the application of educational policies as an instrument designed to promote changes in higher education. The second was to review some models designed to assess student outcomes in higher education. Additionally, institutional stratification and

cultural capital theories were reviewed to enrich the analysis of multiple dimensions involved in the assessment of student outcomes in higher education.

The review of the literature on quality systems provides important information necessary to develop a better understanding of the evaluation policy implemented in Brazil to evaluate the quality of undergraduate programs. The literature regarding purpose of the quality effort addresses important components of the policy formulation, such as the pressures for implementing educational policies, the range of purposes of educational policies, and the problems that policymakers face in promoting policies that address multiple purposes. The different perspectives for defining quality in higher education provide the guidance necessary to examine and criticize how quality is conceptualized in the Brazilian evaluation system. Furthermore, the literature review on different approaches used in educational policies to address quality allows this investigation to develop a more comprehensive understanding of the evaluation instruments, as well as the evaluation outcomes, used in Brazil.

The literature about models for monitoring outcomes in higher education offers a useful framework for assessing the outcomes of the Brazilian evaluation system. The models reviewed stress the importance of taking into consideration input characteristics, environmental aspects, and contextual aspects when analyzing educational outcomes. Furthermore, the discussion about the assessment of educational outcomes in higher education calls attention to the discussion of the use of standardized tests as a valid procedure to measure outcomes. This literature provides the framework for understanding the challenges of the process of assessing psychological outcomes and for identifying

common biases involved in the application of standardized tests. Furthermore, it has been emphasized that caution is necessary in the analysis of the results of standardized tests.

Lastly, institutional stratification and cultural capital provide the theoretical ground that supports some of the claims found in the literature about quality in higher education and models for monitoring outcomes in higher education. Institutional stratification theory provides a framework for understanding the mechanisms by which resources, prestige, and privileges are allocated across types of institutions and the potential impact of the evaluation system and its outcomes on the whole higher education system. Cultural capital provides a different perspective with which to understand differences in educational outcomes by including a critical analysis of the influence of the socioeconomic origins of students on educational outcomes.

CHAPTER III

METHODS

Introduction

This chapter describes the rationale for applying a quantitative approach to achieve the goals of this investigation. The description of the multiple sources of the data is presented, followed by the presentation of the criteria used to select academic fields, student cohort data, undergraduate programs, and undergraduate senior students. This chapter also presents a description of the variables included in the statistical analyses. Furthermore, the analysis procedure used to select fields, programs, and students is described, followed by the discussion of the analytical strategies used to address the research questions posited.

Research Methods Rationale

This investigation employed a quantitative approach to examine the outcomes of the national evaluation system implemented in the Brazilian Higher Education system. Specifically, a quantitative approach offers the opportunity to explore the national data that include student test scores in some academic fields, as well as the outcomes of the evaluation of undergraduate programs in the following dimensions: faculty, curriculum, and facilities.

The literature reviewed argues for exploring the relationship between educational inputs, college environment, and student outputs (Astin, 1993; Burke & Minassians,

2002; Frazer, 1992; Harvey & Green, 1993; Vroeijenstinjn, 1992). The research on student performance and student outcomes frequently explores how they are related to individual characteristics, high school experiences, and college experiences, as well as some institutional characteristics. Drawing from the literature reviewed, this investigation investigated the degree to which student test scores were associated with some aspects of undergraduate programs assessed by ACO, such as faculty, curriculum, and facilities. In addition, some available measures of student background characteristics and institutional characteristics were included in the analyses to control for individual characteristics, as well as institutional characteristics of the college attended.

In order to explore the relationship between educational conditions assessed by ACO and student test scores on the ENC, four important decisions were taken. First, four out of thirteen academic fields were selected (Administration, Law, Civil Engineering, and Dentistry). This decision to narrow the scope of the analysis was taken to offer the possibility of conducting a more detailed analysis of each field, as well as of comparing the similarities and differences between them. Second, one cohort of the program evaluation and one cohort of the student test score were selected, aiming to rule out possible differences in the evaluation procedures carried out in different years, which would make the analysis of the results more complex. Third, a lag of three years was defined between the year the ACO was carried out and the year students took the ENC exam, for reasons that will be explained in the selection criterion of the data, which will be presented later in this chapter.

Data Sources

The data utilized for this investigation came from three different databases administered by two federal agencies of the Ministry of Education (MEC): 1) the student exit exam (ENC) database; 2) the evaluation of educational conditions (ACO) database; and 3) the national information system for higher education (SIEdSup).

The first national database contains information about the application of the national exit exam (ENC). This database provided all the information related to students (age, race, SES background, student involvement in college, admission cohort, and test score) and some of the information related to undergraduate programs (e.g., mean test score and program selectivity). The logistics involved in applying the national exit test (ENC) are coordinated by INEP (National Institute for Educational Studies and Research, a federal agency). Every year, a national committee for each academic area is appointed by the federal government through the nomination of 7 specialists selected from a list of names compiled by Council of Presidents of Brazilian Universities (CRUB), professional associations, SESU, among others (Instituto Nacional de Estudos e Pesquisas Educacionais [INEP], 2001). This committee is responsible for defining the content domain that guides the preparation of the test. The INEP then contracts an external organization to design the test, administer the test, process the results, and tabulate the responses of the SES questionnaire and survey. The test, one for each academic field, is designed considering the content domain defined by the national committee, which includes multiple-choice questions as well as discursive questions. The exception was the test applied for Civil Engineering, which had only open-format questions. The test is then

applied once a year, on the same day, to all students registered by their institutions. On the day of the test, students are asked to fill out the SES questionnaire and survey before answering the test. The responses of this questionnaire, as well as the test scores, are confidential and only those students who agree to receive a report by mail are allowed to know their test score.

The second database contains the results of the program evaluation carried out once every five years to assess the educational conditions of undergraduate programs (ACO). This program evaluation is coordinated by the Secretariat of Higher Education (SESu), another federal agency of the Ministry of Education (MEC), and includes the assessment of the following educational dimensions of each undergraduate program, referred hereafter as educational conditions: faculty, curriculum, and facilities. In the year the program evaluation is to occur, the SESu publishes the list of undergraduate programs to be evaluated. Then, a manual containing the assessment procedures is sent to all institutions whose programs are going be evaluated. This manual presents all the information and documents that program administrators have to make available on the date the specialists visit the institution, as well as the procedures used to determine the assessment outcome in regard to each educational dimension (faculty, curriculum, and facilities). The final evaluation rating for each dimension is measured in a four-point scale, whose information ranges from "insufficient" conditions, "regular" conditions, "good" conditions, and "very good" conditions.

The third database is the national information system for higher education (SIEdSup) administered by the Secretariat of Higher Education (SESu) of the MEC. At

the beginning of each academic year (January in Brazil), all higher education institutions are required to complete a census that collects information about institutional characteristics, undergraduate programs, faculty, staff, finance, and facilities. The data from this national database includes information about institutional characteristics (e.g., institutional type and administrative control), as well as geographical information (state, and region).

Selection Criteria for Academic Fields

The scope of this investigation was narrowed to undergraduate programs and undergraduate students in the following academic fields: Administration, Law, Civil Engineering, and Dentistry. Despite the fact that the federal government made available all data from undergraduate programs assessed by ACO between 1997 and 2000 in 13 academic fields, the decision to select four of them was taken in order to reduce the scope of the analysis and make it possible to conduct a more careful analysis of the data. Several aspects were considered in selecting these academic fields.

First, Administration and Law were selected for being the biggest fields in Brazilian Higher Education in terms of the number of programs and the number of students enrolled. This criterion allowed for including programs with considerable diversity in terms of institutional characteristics as well as student characteristics.

Second, all academic fields that were also evaluated by ACO in the same period of the evaluation carried out in Administration and Law (which was between 1997 and 1998) were considered. This criterion narrowed the number of academic fields from thirteen to six fields: Civil Engineering, chemical engineering, veterinary medicine, and

Dentistry. This decision was taken to ensure that the general procedures, as well as the structural conditions of Brazilian Higher Education, were the same at the time the program evaluation was carried out. Therefore, all academic fields examined in this investigation had their programs evaluated by ACO between 1997 and 1998.

Third, academic fields with a small number of programs were excluded from the analysis. Therefore, chemical engineering and veterinary medicine were excluded because the former had only 43 programs and the latter only 37 programs assessed in the whole country. The criteria used led to the selection of fields with very interesting characteristics: Administration and Law (the largest fields), Civil Engineering (technological field), and Dentistry (a very selective field).

Selection Criteria for the Student Cohort Data

Another important aspect of the research design is related to the selection of the student cohort data to be used for investigating the research question of this investigation. Despite the fact that the data available for this investigation included the cohorts in the period between 1997 and 2002, only one cohort was used. The decision to select the 2001 cohort was based on the following reasons.

The first reason relates to the time period between the year the educational conditions of the programs were evaluated by ACO and the year students took the ENC. It is important to recall that only students in their last academic year can participate in the ENC. Because the primary objective of this investigation was to explore how certain educational characteristics assessed by the ACO were related to student performance at the ENC, it was necessary to select an ENC cohort whose students had experienced the

conditions assessed by the ACO during the time they were attending college. This means that if the results of the ACO assessment reflect the educational conditions assessed in 1997/1998, most of the students should have been attending college during this time. The percentage of students who participated in the ENC in each year by admission cohort is displayed in Table 3.1. The 2001 student cohort was selected because most of the students were admitted between 1997 and 1999; therefore, they had the opportunity to experience the conditions evaluated by ACO between 1997 and 1998.

Table 3.1 - Percentage of Students who Took the ENC by Admission Cohort

Academic field	ENC year	Admission Cohort						
		Before 1996	1996	1997	1998	1999	2000	1997-1999
Admistration	1999	52.5%	45.1%	1.1%	0.4%	0.3%	0.0%	1.7%
	2000	21.8%	33.0%	43.5%	1.0%	0.3%	0.2%	44.8%
	2001	11.0%	11.1%	29.1%	46.9%	1.2%	0.4%	77.2%
	2002	5.3%	4.1%	9.3%	29.2%	49.9%	1.2%	88.4%
Law	1999	94.6%	4.7%	0.4%	0.1%	0.1%	0.0%	0.5%
	2000	31.0%	66.4%	1.7%	0.4%	0.2%	0.0%	2.3%
	2001	8.4%	23.9%	65.1%	1.4%	0.6%	0.2%	67.1%
	2002	3.8%	5.3%	27.4%	60.8%	1.6%	0.5%	89.9%
Civil Engineering	1999	97.4%	1.4%	0.4%	0.2%	0.1%	0.0%	0.7%
	2000	66.2%	30.6%	1.2%	1.2%	0.4%	0.0%	2.7%
	2001	31.5%	32.6%	33.0%	1.9%	0.6%	0.2%	35.6%
	2002	17.3%	14.1%	32.7%	32.1%	2.2%	1.1%	67.0%
Dentristry	1999	64.1%	35.5%	0.3%	0.0%	0.0%	0.0%	0.3%
	2000	13.5%	53.4%	32.7%	0.2%	0.1%	0.0%	32.9%
	2001	2.7%	9.9%	56.2%	30.7%	0.3%	0.1%	87.2%
	2002	1.0%	1.5%	13.5%	53.4%	30.2%	0.3%	97.1%

One might argue that educational conditions may have changed between the time the ACO assessment was carried out in 1997/1998 and the time that the 2001 cohort took the test. Such changes could make the interpretation of the results of this investigation somehow biased. However, this assessment is carried out once every five years and program administrators have prior knowledge of when their programs will be evaluated in the next assessment cycle. The knowledge of the assessment period (which was between 1997 and 1998), as well as the criteria used to assess educational conditions, would lead to the development of some institutional actions preceding the assessment in order to comply with some standards and, therefore, to achieve better assessment outcomes. Therefore, it can be assumed that most of the changes promoted by institutions to make the educational conditions of their undergraduate program better had happened prior to the realization of the 1997/1998 ACO. Besides, it is important to recognize that higher education institutions are organizations that do not promote radical changes in a short period of time, due to the kind of institutional culture as well as the institutional complexity. This makes the selection of the 2001 results a reasonable cohort in which to examine the effects of the educational conditions assessed by ACO only a few years earlier in 1997/1998.

Selection Criteria for Undergraduate Programs

The criteria used to select undergraduate programs were as follows. All undergraduate programs from Administration, Law, Civil Engineering, and Dentistry were initially selected. Because the primary objective of this investigation is to explore the relationship between the outcomes of the ENC and ACO, only those undergraduate

programs that had data from both assessments were considered. Furthermore, only undergraduate programs that received one of the four valid assessment ratings (e.g., "insufficient", "regular", "good", and "very good") were selected. This criterion excluded two programs from Administration, one from Law, and two from Dentistry whose ACO assessment rating was "without concept" (SC) because the evaluation was not carried out.

Table 3.2 presents the number of undergraduate programs assessed by ACO in 1998 and by ENC in 2001. As can be seen, more than two-thirds of the programs from Administration and Law and more than four-fifths of the programs from Civil Engineering and Dentistry were evaluated by both instruments, which indicates that the sample of programs included in this investigation is likely to be a highly representative sample of programs evaluated in 2001. It is also important to point out that the majority of programs excluded from this investigation were due to the lack of ACO assessment outcomes carried out in 1997/1998. Because only those programs whose students participated in the ENC in 1996 and 1997 were assessed by the ACO in 1997/1998, it is assumed that most of the programs excluded were recently created and did not have a cohort ready to graduate by the period the program evaluation was carried out.

Table 3.2 - Number of Undergraduate Programs Assessed by ACO in 1998 and ENC in 2001

Academic Fields	Programs Assessed by ACO (1998)	Programs Assessed by ENC (2001)	Programs Assessed by ACO (1998) and ENC (2001)	% Programs assessed by ACO (1998) and ENC (2001)
Administration	336	488	330	68%
Law	192	271	189	70%
Civil Engineering	104	123	103	84%
Dentistry	84	102	82	80%

Source: Database from SESu/MEC and INEP

Selection Criteria for Undergraduate Senior Students

Initially, all senior students from Administration, Law, Civil Engineering, and Dentistry who participated in the 2001 ENC were selected. Then, students from undergraduate programs evaluated by ACO and ENC were considered. Following that, only students whose test score were included as valid cases for the 2001 final report published by INEP were included in this investigation. This selection was based on two dummy variables provided by INEP, indicating a valid presence in the exam (TP_PRES) and the first time a given student took this exam (IN_PRIM).

After applying these criteria, all students with zero test scores were excluded for the following reason. The legislation that created this exam established that all students must register and be present on the date the test is administered. Because the test score does not affect the graduation process, some students might decide to turn in a blank test. Although this practice was frequently experienced initially as a form of student protest against this evaluation procedure, as can be seen in Table 3.3, the percentage of blank tests turned in has been dramatically reduced in recent years.

Table 3.3 - Percentage of Blank Tests Turned in by Year of Realization of ENC

Academic Field	Year of ENC					
	1996	1997	1998	1999	2000	2001
Administration	8.4%	0.4%	0.4%	0.2%	0.2%	0.2%
Law	11.4%	1.9%	1.4%	0.4%	0.5%	0.4%
Civil Engineering	32.1%	13.9%	6.5%	6.8%	5.9%	1.2%
Dentistry		1.0%	0.2%	0.1%	0.1%	0.4%

Source: DAES/INEP - ENC/2001

Final Sample of Undergraduate Programs and Undergraduate Students

After applying the criteria for selecting academic fields, undergraduate programs, and undergraduate students, the final sample comprises 698 undergraduate programs and 77,085 undergraduate students from Administration, Law, Civil Engineering, and Dentistry. Breaking down the number of programs and students by academic field, the final sample included 327 programs and 30,776 students in Administration, 189 programs and 34,834 students in Law, 102 programs and 4,880 students in Civil Engineering, and 80 programs and 6,595 students in Dentistry, as presented in Table 3.4. This sample represents more than two-thirds of all students who took the ENC in 2001 and over 85% of all students from programs assessed by ACO in 1998 and ENC in 2001. The lower percentages can be found in Administration and Law due to the strong expansion of undergraduate programs since 1995---this resulted in many programs being excluded from the ACO evaluation in 1997/1998.

Table 3.4 - Number of Students in Programs Assessed by ACO (1998) and ENC (2001)

Academic Fields	Students in Programs Assessed by ENC (2001)	Students in Programs Assessed by ACO (1998) and ENC (2001)						
		Total		Response Rate to SES questionnaire and survey		Final Sample		
	(a)	(b)	(c)=(b/a)	(d)	(e=d/b)	(f)	(g)=(f/a)	(h)=(f/b)
Administration	45,578	36,293	80%	33,472	92%	30,776	68%	85%
Law	50,144	40,719	81%	37,801	93%	34,834	69%	86%
Civil Engineering	5,977	5,546	93%	5,192	94%	4,880	82%	88%
Dentistry	8,945	7,452	83%	7,040	94%	6,595	74%	88%
Total	110,644	90,010	81%	83,505	93%	77,085	70%	86%

Dependent variable

The dependent variable for this investigation is the standardized test score from the exit exam (ENC). For each student, a raw test score is determined by calculating the average score of the multiple-choice questions and the discursive questions. For Civil Engineering, however, the raw test score applied only to the open-format questions because the exam includes only this type of question. Descriptive analysis of the student test score by academic field is provided in Table 3.5, revealing different mean test scores and standard deviations across fields. The observed differences are due to the fact that each academic field had a different exam designed to assess knowledge and skills unique for the field. The highest mean test score was found in Dentistry, followed by Law, Administration, and Civil Engineering. In regard to the standard deviation, Dentistry presented the smallest standard deviation among the fields included in this investigation while Civil Engineering had the highest standard deviation, which might be related to the fact that the test was designed with only open questions.

Table 3.5 - Mean and Standard Deviation of Exit Exams (ENC), by Academic Field

Academic Field	N	Mean	SD	Minimum	Maximum
Administration	30,776	31.90	11.14	1.30	76.30
Law	34,834	40.61	13.88	1.25	90.00
Civil Engineering	4,880	29.08	15.92	1.00	92.00
Dentistry	6,595	54.03	9.25	17.50	87.50

Another important characteristic of this exam relates to the distribution of the test scores around the mean. This information is important because the ENC is a standardized exam known as a norm reference test (NRT) whose main characteristic is to provide the relative position of a given student in relation to all students taking this exam. Therefore, the expected outcome of this type of exam should follow a normal distribution if the test achieves its goal. As presented in Figure 3, the distribution of the test scores approached a normal distribution in all fields examined.

Figure 3 – Distribution of Test Scores by Academic Field

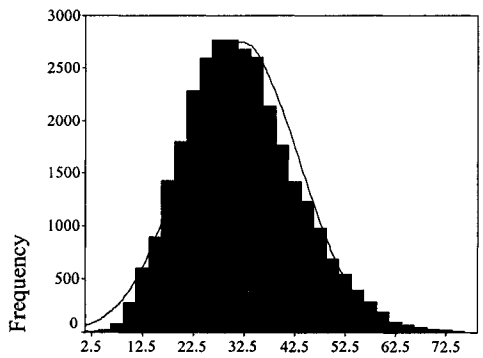

Figure 3.1 - Test Score - Administration

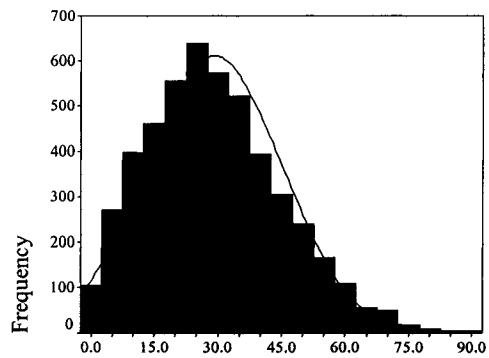

Figure 3.3 - Test Score - Civil Engineering

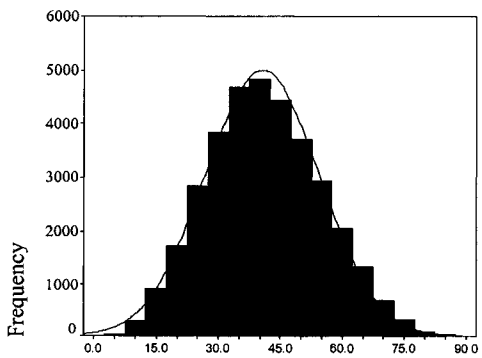

Figure 3.2 - Test Score - Law

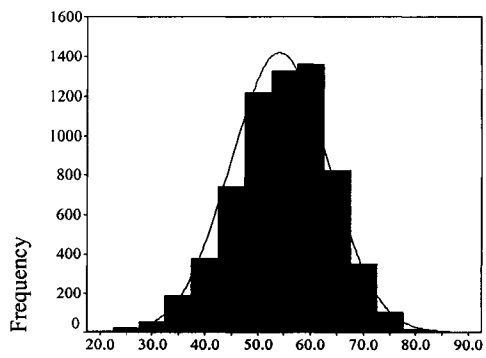

Figure 3.4 - Test Score - Dentistry

In order to make the results of the regression analysis easier to examine, a linear transformation of the test score was conducted. The linear transformation followed the same transformation procedure used to define the standardized mean of undergraduate programs (INEP, 2001). Martuza (1977) defines the linear transformation used in this investigation as the College Entrance Examination Board (CEEB) transformation, which has the following form:

$$Z' = 100 * Z + 500$$

where Z' represents the student standardized score, Z represents the normalized test score, 500.0 is the mean of the new distribution, and 100.0 is the standard deviation of the new distribution.

This transformation brought some advantages for the analyses performed. Martuza (1977) points out that this transformation process has some advantages over the use of Z scores because all scores are reported in a positive way. In addition, it minimizes the loss of information due to the rounding process, without changing the shape of the original distribution. Furthermore, the use of this new scale facilitated the interpretation of results over the use of raw test scores. Because all the information produced by this type of test (NRT) is relative to the performance of others, the regression coefficients were analyzed in relation to the number of standard deviations of the new distribution, which was 100.0. For instance, a coefficient of 50.0 is interpreted as being associated with a half standard deviation movement toward to the right tail of the distribution, while a negative coefficient is associated with a movement to the left tail. This new scale also made it possible to compare the regression results among academic fields, since

standardized scores from all four academic fields present the same mean and standard deviation.

Independent variables

Even though the review of the literature pertinent to this research refers to a vast list of variables that could affect student performance, the decision to use data available at the national level limited the selection of the independent variables to the information available in the ENC, ACO, and SIEdSup databases. This section presents the descriptive analysis of the independent variables at the institution level, program level, and student level included in this investigation.

Institution-Level Variables

The institution-level data were included in the analyses to explore the effects of some institutional characteristics on student test scores, such as institutional type and administrative control (see Appendix A for definition of types of institutions and administrative control). Administrative control was coded in a series of five dummy variables indicating one of the following types of institutional control: a federal-controlled, a state-controlled, a local-controlled, a religious/community-controlled, or a for-profit institution. Institutional type was coded into three dummy variables indicating whether the institution is a university, a university-center, or college. The number and the percentage of undergraduate programs included in this investigation are shown in Table 3.6 according to the administrative control and institutional type. In addition, a third

Table 3.6 - Number of Programs by Academic Field, Administrative Control, and Institutional Type

Institutional Characteristics	Adminis-tration		Law		Civil Engineering		Dentistry		Total	
All Institutions	327	100%	189	100%	102	100%	80	100%	698	100%
Public x Private										
Private	245	75%	135	71%	52	51%	40	50%	472	68%
Public	82	25%	54	29%	50	49%	40	50%	226	32%
Administrative Control										
Federal	36	11%	35	19%	32	31%	27	34%	130	19%
State	30	9%	11	6%	15	15%	12	15%	68	10%
Local	16	5%	8	4%	3	3%	1	1%	28	4%
Religious/Community	115	35%	76	40%	33	32%	24	30%	248	36%
Private for Profit	130	40%	59	31%	19	19%	16	20%	224	32%
Institutional Type										
University	164	50%	121	64%	82	80%	70	88%	437	63%
University Center	40	12%	16	8%	8	8%	3	4%	67	10%
College	123	38%	52	28%	12	12%	7	9%	194	28%

aggregation was also provided to illustrate the presence of private and public institutions as providers of education in these academic fields.

Three important findings emerged from the information provided in this table. The first relates to the relationship between the cost of education and the public provision of programs. As can be seen from this table, public institutions were responsible for half of all undergraduate programs in the most expensive fields (Civil Engineering and Dentistry) against approximately one-fourth of the undergraduate programs in the least expensive fields (Administration and Law). This finding also draws attention to the importance of the federal government as a provider of nearly one-third of undergraduate programs in Civil Engineering and Dentistry. Unlike U.S. higher education, state and

local governments play a different role in providing higher education; these institutions tend to be created by state and local government from poor regions, especially in the Northeast, and usually offer less expensive undergraduate programs. The exception is São Paulo State, which maintains three large, high-quality research universities, even though these universities were responsible for less than ten percent of the total enrollment in this state in 2001.

The second finding is related to the strong presence of private institutions in Brazilian Higher Education. Private institutions—community- or religious-controlled institutions and for-profit institutions—were responsible for two-thirds of all programs included in this research. Private, for-profit institutions were responsible for nearly forty percent of programs from Administration, while private, non-profit institutions (community- or religious-controlled institutions) were responsible for between thirty and forty percent of the programs included in this investigation. Indeed, the participation of private institutions in these academic fields has remarkably increased in recent years due to the expansion of the undergraduate education in Brazil (see Table 1.1).

The third finding relates to the predominance of the university model of higher education in Brazil. Data from Table 3.6 show that universities maintained over four-fifths of all undergraduate programs from Civil Engineering and Dentistry and more than half of all programs from Administration and Law. It also important to point out that the recent expansion of undergraduate education in Brazil in these fields has been promoted largely in universities and university-center institutions. The exception is Administration,

where the number of programs provided by colleges and assessed by ENC in 2001 increased by 78 percent.

The results of the cross-tabulation analysis presented in Table 3.7 provide relevant information to better understand the relationship between institutional type and administrative control in the fields examined. Of all institutions that maintain undergraduate programs in these fields, nearly all federal and state institutions and more than two thirds of non-profit institutions are universities. The percentage of federal and state universities is larger than ninety percent in these fields; the only exception is in Administration, where 80 percent of state institutions are universities. Local governments and for-profit corporations tend to maintain smaller institutions (mostly colleges) and to offer Administration and Law programs (the less expensive ones). Another interesting aspect is related to the presence of for-profit organizations in the Dentistry field. Due to the relatively high cost of Dentistry, three-quarters of for-profit institutions offering these programs are universities, which reinforce the assumption that high-cost programs are mostly provided by either public institutions or private universities.

In summary, there are key differences between public and private institutions. On one hand, public institutions can be characterized by the following aspects: 1) the vast majority are universities, 2) funded by public appropriations, 3) do not charge tuitions, 4) have a near-monopoly of research (especially federal and some state institutions), and 5) a strong presence in graduate education. On the other hand, private institutions can be described by having the following characteristics: 1) most non-profit institutions are universities and most for-profit institutions are university-centers or colleges, 2) funded

Table 3.7 - Cross-tabulation of Administrative Control and Institutional Type

Administrative Control	Intitutional Type									Total		
	University			University Center			College					
	N	%[a]	%[b]	N	%[a]	%[b]	N	%[a]	%[b]	N	%[a]	%[b]
Administration												
Federal	34	21%	94%	-	-	-	2	2%	6%	36	11%	100%
State	24	15%	80%	-	-	-	6	5%	20%	30	9%	100%
Local	2	1%	13%	1	3%	6%	13	11%	81%	16	5%	100%
Relig./Community	78	48%	68%	10	25%	9%	27	22%	23%	115	35%	100%
Private for Profit	26	16%	20%	29	73%	22%	75	61%	58%	130	40%	100%
Total	164	100%	50%	40	100%	12%	123	100%	38%	327	100%	100%
Law												
Federal	35	29%	100%	-	-	-	-	-	-	35	19%	100%
State	10	8%	91%	-	-	-	1	2%	9%	11	6%	100%
Local	2	2%	25%	-	-	-	6	12%	75%	8	4%	100%
Relig./Community	56	46%	74%	6	38%	8%	14	27%	18%	76	40%	100%
Private for Profit	18	15%	31%	10	63%	17%	31	60%	53%	59	31%	100%
Total	121	100%	64%	16	100%	8%	52	100%	28%	189	100%	100%
Civil Engineering												
Federal	30	37%	94%	-	-	-	2	17%	6%	32	31%	100%
State	15	18%	100%	-	-	-	-	-	-	15	15%	100%
Local	2	2%	67%	-	-	-	1	8%	33%	3	3%	100%
Relig./Community	27	33%	82%	1	13%	3%	5	42%	15%	33	32%	100%
Private for Profit	8	10%	42%	7	88%	37%	4	33%	21%	19	19%	100%
Total	82	100%	80%	8	100%	8%	12	100%	12%	102	100%	100%
Dentistry												
Federal	25	36%	93%	1	33%	4%	1	14%	4%	27	34%	100%
State	12	17%	100%	-	-	-	-	-	-	12	15%	100%
Local	1	1%	100%	-	-	-	-	-	-	1	1%	100%
Relig./Community	20	29%	83%	1	33%	4%	3	43%	13%	24	30%	100%
Private for Profit	12	17%	75%	1	33%	6%	3	43%	19%	16	20%	100%
Total	70	100%	88%	3	100%	4%	7	100%	9%	80	100%	100%

a - Percentage related to the total number of programs in each of type of institution (by column)

b - Percentage related to the total number of programs in each type of adm. control (by line)

predominantly by tuition and fees, 3) fewer resources available for conducting research (due to strong restrictions on allocating public resources to private institutions), 4) offer less expensive programs, and 5) enroll more than two-thirds of undergraduate students.

Program-Level Variables

Two types of information about programs were included in the analysis. The first type relates to the results of the ACO assessment carried out between 1997 and 1998. The second type relates to the selectivity of the admission process. For each type of program characteristics, a brief presentation of each variable is presented, followed by the descriptive analysis.

Evaluation of Educational Conditions (ACO). The results of the ACO assessment carried out between 1997 and 1998 provide the information necessary to explore the degree to which the outcomes of the ACO assessment are associated with student test scores. Each undergraduate program assessed by ACO receives an evaluation score that represents the result of the evaluation carried out by specialists (ACO) in each of the following dimensions: faculty, curriculum, and facility (educational conditions).

The criteria used to evaluate each of these dimensions varied across programs. For each field, a national commission was created to define the procedures applied to assess undergraduate programs, as well as to set the standards for each of these dimensions. For faculty, the aspects evaluated included the following: faculty credentials (B.A., specialization, M.A., Ph.D.), dedication to the program (full-time vs. part-time), publications, career plans, and student-faculty ratio, among others. For curriculum, the

aspects evaluated included the following: curriculum organization, research and publications, and student-class ratio, among others. For facilities, the aspects evaluated included the following: libraries, equipment, laboratories, classrooms, and faculty offices. Despite the different criteria used in these fields, the final outcome for each of these dimensions is a four-point scale representing the degree to which a program has not achieved the minimum standards ("insufficient" condition), has achieved the minimum standards ("regular" condition), has exceed the minimum standards ("good" condition), or has achieved a high standard of excellence ("very good" condition),

Table 3.8 displays the number and the percentage of programs by academic field according to the results of the ACO assessment. The data reveal that at least two-thirds of the programs received a "regular" or "good" assessment outcome in each dimension assessed. The exceptions were Administration, where eighty-eight percent of all programs received "regular" or "good" evaluations of faculty, and Law, where half of the programs received "regular" and "good" evaluation outcomes in curriculum and facility.

With regard to the faculty evaluation ratings, the highest percentage of programs with "good" and "very good" evaluation ratings was found in Dentistry, while the highest percentage of programs with "insufficient" evaluation ratings was found in Law. The data related to the curriculum evaluation ratings show a considerable difference in the percentage of programs evaluated as "insufficient", where the percentage of programs from Law was quite a lot higher than from the other fields, as seen in the results by academic field. The results regarding the facility evaluation ratings revealed that a higher percentage of programs received "good" and "very good" evaluation ratings in the

Table 3.8 - Results of the ACO Assessment for Faculty, Curriculum, and Facility, by Academic Field

Educational Conditions	Administration		Law		Civil Engineering		Dentistry	
Faculty (ACO)								
Insufficient	19	6%	52	28%	15	15%	2	3%
Regular	166	51%	63	33%	40	39%	15	19%
Good	122	37%	59	31%	31	30%	41	51%
Very Good	20	6%	15	8%	16	16%	22	28%
Total	327	100%	189	100%	102	100%	80	100%
Curriculum (ACO)								
Insufficient	56	17%	60	32%	2	2%	12	15%
Regular	111	34%	47	25%	26	25%	24	30%
Good	112	34%	45	24%	50	49%	27	34%
Very Good	48	15%	37	20%	24	24%	17	21%
Total	327	100%	189	100%	102	100%	80	100%
Infrastructure (ACO)								
Insufficient	28	9%	47	25%	6	6%	11	14%
Regular	75	23%	26	14%	20	20%	34	43%
Good	129	39%	77	41%	54	53%	21	26%
Very Good	95	29%	39	21%	22	22%	14	18%
Total	327	100%	189	100%	102	100%	80	100%

facility dimension than in faculty and curriculum. Furthermore, the highest proportion of programs with "insufficient" evaluation ratings in all program dimensions evaluated was found in Law.

Program Selectivity. A variable representing the program selectivity level was created based on the ratio of the number of applicants to the number of spaces for freshmen gathered from the ENC database. Due to the fact that there is no other measure available nationwide that could be used as a proxy for pre-college preparation (such as SAT and ACT), selection ratio for undergraduate programs measured in terms of the number of students applied and the number of freshmen enrolled was included as a proxy for program selectivity. The basic assumption is that programs with a high selectivity ratio would be able to select better students than programs with a low selectivity ratio.

The operational concept used to create this new variable was the following. For each undergraduate program, selectivity was calculated based on the mean selectivity from 1997 to 2001. To avoid excluding some cases due to lack of information necessary to create this variable, missing data were replaced by the average of program selectivity from other programs in the same state. The number of cases with missing data is displayed in Table 3.9. The decision to replace missing data by the mean selectivity of the state was taken because the data showed that the selectivity of programs varied a lot from state to state. This variation might be related to structural characteristics of higher education in each state, which certainly is influenced by the accessibility of higher education offered by colleges and universities. Therefore, it is necessary to acknowledge that the program selectivity methodology might have possible measurement problems due to the fact that the demand for higher education varies across states.

Table 3.9 - Number of Programs with Missing Information about Selectivity

Academic Fields	Programs with Valid Information about Selectivity		Programs with Missing Information about Selectivity		Total Number of Programs Included in this Investigation
Administration	324	99%	3	1%	327
Law	187	99%	2	1%	189
Civil Engineering	86	84%	16	16%	102
Dentistry	80	100%	0	0%	80

The selectivity level in each academic field is displayed in Table 3.10, according to the institutional characteristics. As presented in the table, public institutions had a higher selectivity ratio as well as a higher standard deviation than private institutions in

all fields, which is a result of the selectivity presented by federal and state institutions. This means that the selectivity among public institutions varied more than the selectivity among private institutions. With regard to the institutional type, universities were much more selective than university-centers and colleges in all fields. One important aspect relates to the fact that the difference in the selectivity level is more expressive when analyzed by the type of administrative control than by the institutional type. This finding might suggest that the free-tuition policy promoted by public institutions might be playing a significant role in the admissions process.

Table 3.10 - Program Selectivity by Institutional Characteristics

Institutional Characteristics	Administration		Law		Civil Engineering		Dentistry	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All Institutions	4.27	4.07	10.74	8.23	4.56	3.86	16.50	10.11
Public x Private								
Private	2.71	1.72	6.97	4.25	2.06	1.48	10.76	6.04
Public	9.02	5.33	19.68	8.51	7.35	3.79	23.69	9.60
Administrative Control								
Federal	10.74	3.50	22.95	8.06	7.23	3.15	21.70	7.47
State	10.75	6.22	17.29	3.86	9.55	4.65	34.28	10.56
Local	3.42	1.85	10.02	6.09	1.93	0.76	9.88	.
Religious/Community	2.69	1.67	6.79	4.26	2.24	1.72	11.73	6.55
Private for Profit	2.73	1.76	7.16	4.27	1.79	1.00	9.31	5.02
Institutional Type								
University	5.88	5.24	12.72	9.52	5.09	3.67	17.03	10.36
University Center	2.90	1.70	7.55	3.97	1.79	1.32	14.46	11.36
College	3.09	2.37	7.74	4.36	3.09	5.18	12.73	7.32

Student-Level Variables

After analyzing the availability of data, the following variables were included in this investigation: defining characteristics (age and race), socioeconomic background (SES), student involvement in college, and admission cohort. Due to the fact that all student-level variables were collected from an existing survey instrument administered annually by the federal government, some background variables that would be desirable were not included, such as gender and an equivalent to the ACT/SAT test score. Information about gender was requested by the researcher but was not provided. In regard to a measure of pre-college academic achievement, there is no measure available at the national level that can be included in this investigation to control for previous academic performance and abilities. Unlike SAT/ACT scores, which are widely used in the U.S. college admissions process, the admissions process in Brazilian colleges and universities is based on entrance exams known as *vestibular*, designed and administered by each institution or group of institutions. It means that most of the institutions have student test scores measured by their *vestibular* exams. However, the test scores are not equivalent because each institution designs its own entrance exam and because the *vestibular* scores are not available in any national databases.

Student background characteristics included the following: demographic characteristics and socioeconomic status. Demographic characteristics were broken down into a series of dummy variables indicating the age group (up to 24 years old, between 25 and 30 years old, and older than 30 years) and race (White, Black/Mulatto, Asian, and Native). Socioeconomic status was measured through the following variables: family

income, father's education level, and type of high school attended. Family income (measured in Brazilian currency) was broken down into five dummy variables indicating the family income in one of the following intervals: less than R$ 601; between R$ 601 and R$ 2,000; between R$ 2,001 and R$ 4,000; between R$ 4,001 and R$ 10,000; and R$ 10,001 and greater. These income intervals were arbitrarily defined in the survey instrument (SES socioeconomic questionnaire and survey); this means that they do not have any direct link with the income distribution of the Brazilian population as discussed later in this chapter. The original variable indicating the highest level of education attained by student's father was measured using a five-point scale ranging from 0 to 4 (0 for less than elementary education, 1 for elementary education, 2 for middle education, 3 for high school, 4 for college). For the purpose of this investigation, father's educational level was defined by three dummy variables indicating having less than a high school degree, having a high school degree, and having at least a college degree. With regard to the type of high school attended, three dummy variables were created to indicate whether students attended only public high schools, only private high schools, or some private and some public high schools.

The other important construct derived from some questions of the SES questionnaire and survey was one that indicated the degree to which individual students were involved in some activities in college. Student involvement was created based on a six-point scale ranging from zero to five, depending on the degree to which students were involved in five types of activities. Students were asked to answer five questions that described their opportunity to be involved in some activities (e.g., going to the library,

looking for out-classroom orientation, and participating in extra academic activities, extra-curricular activities, and events). For each question, one point was assigned when students said they were involved in one of the activities asked and zero points were assigned if they were not involved in any activities. Then, the student involvement score was defined by the sum of the values assigned for each of the five questions.

The last set of data related to the admissions cohort. Data about the year each student was admitted to college provided very interesting information, which was used in this research to explore the relationship between time taken to graduate and student performance on the ENC. Because only senior students in their last academic year should participate in this exam as a requirement for requesting the diploma, the difference between the year they took the ENC (2001 in this investigation) and the year they were admitted to college could be used as a measure of time to graduate. The data regarding the admissions cohort were broken down into four dummy variables indicating the following: being admitted in 1995 or before, being admitted in 1996, being admitted in 1997, and being admitted in 1998 or later.

Descriptive statistics of the student-level data, shown in Table 3.11, provide basic information about the average profile of undergraduate senior students included in this investigation. This table displays the mean of all independent variables for each academic field, where separate analyses were carried out within each field. Complete descriptive statistics such as mean, standard deviation, minimum, and maximum values of these variables can be found in Appendix B (see Appendix B.1.1, Appendix B.1.2, Appendix

Table 3.11 - Mean of Student-Level Variables, by Academic Field

Variables	Adminis-tration	Law	Civil Eng.	Dentistry	Total
Student-Level Variables (N)	30,776	34,834	4,880	6,595	77,085
Admission Cohort					
Admitted in 1995 or Before	0.11	0.08	0.32	0.03	0.11
Admitted in 1996	0.12	0.23	0.33	0.11	0.18
Admitted in 1997	0.31	0.66	0.32	0.57	0.49
Admitted in 1998 or Later	0.46	0.02	0.03	0.29	0.22
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	0.53	0.56	0.60	0.82	0.57
Age: 25-30 Years Old	0.29	0.24	0.32	0.17	0.26
Age: Older than 30 Years Old	0.18	0.20	0.08	0.02	0.17
Race					
White	0.83	0.84	0.81	0.85	0.83
Black/Mulato	0.13	0.13	0.15	0.09	0.13
Asian	0.03	0.02	0.03	0.05	0.03
Native	0.01	0.01	0.01	0.00	0.01
SES Related Variables					
Family Income					
Income: <= R$ 600	0.04	0.04	0.05	0.02	0.04
Income: R$ 601 to R$ 2.000	0.32	0.28	0.30	0.19	0.29
Income: R$ 2.001 to R$ 4.000	0.34	0.32	0.31	0.36	0.33
Income: R$ 4.001 to R$ 10.000	0.23	0.28	0.27	0.35	0.26
Income: >= R$ 10.001	0.07	0.08	0.08	0.08	0.08
Father Educational Level					
Father Having less than HS Degree	0.49	0.36	0.31	0.21	0.40
Father Having HS Degree	0.23	0.22	0.24	0.23	0.22
Father Having College Degree	0.28	0.42	0.45	0.56	0.38
Type of High Scholl Attended					
All Private HS	0.43	0.53	0.56	0.72	0.51
All Public HS	0.41	0.31	0.31	0.15	0.34
Some Public / Some Private	0.15	0.16	0.13	0.13	0.15
Involvement in College Activities					
Student Involvement	3.69	4.07	4.29	4.84	-

Note: See Appendix B (Table B.1.1, Table B.1.2, Table B.1.3, and Table B.1.4) for complete description of variables

B.1.3, and Appendix B.1.4). A brief comment about student characteristics is presented for each academic field, followed by a comparison of characteristics between fields.

Descriptive analysis of the admission cohort reveals that two-thirds of students from Administration and Law took this test within four to five years of admission to college, compared with eighty-eight percent from Dentistry and thirty-five percent from engineering. The difference between fields might be associated with the program length and/or time to graduate. Regarding program length, the data reveal that Administration and Dentistry are likely to have either four- or five-year programs, which can be seen by the percentage of students graduating in four to five years. In contrast, Law and Civil Engineering are typically five-year programs. Regarding the time to graduate, the data also reveal that two-thirds of students from Civil Engineering took more time to graduate since they took this test at least seven years after initial college admission.

The data presented in Table 3.11 profile the average student who took this test in terms of demographic characteristics and SES background. In general, the vast majority of students were younger than 30 years old at the time of the exam and were predominantly White. Regarding the SES background, two-thirds of students came from families with income greater than to R$ 2.000,00, sixty percent had parents with at least a high school degree, and fifty-one percent attended private high school.

Despite this general description of the average student included in this investigation, some differences exist between fields that should be presented. When compared with other fields, Dentistry had a higher percentage of young students who attended private high school and whose families had superior SES background.

Following the same characterization used to define students from Dentistry, the data show that students from engineering had a higher profile than students from Law, which in turn presented a superior profile to their counterparts from Administration. In addition, students from Dentistry presented a higher involvement in activities during college than those from Civil Engineering, Law, and Administration.

In order to show that the profile of students who took the ENC exam is considerably higher than the average Brazilian population, a comparison of the data from the 2000 census and the data from the socioeconomic (SES) questionnaire and survey is displayed in Table 3.12. The data indicate the class stratification of the Brazilian higher education system. In this instance, Whites are over-represented and Blacks/Mulattos are under-represented in the sample utilized in this investigation. As can result, Whites accounted for 53% of the Brazilian population in 2000 but represent 83% of students from this sample. The differences associated with income levels are even more significant. While sixty-seven percent of the population earned less than R$ 453 in 2000, students from families with income less than R$ 600 accounted for only four percent of the sample. On the other hand, barely eight percent of the population earned more than R$ 1,500 in 2000, but students from families with income higher than R$ 2,000 accounted for sixty-seven percent of the sample. In regard to the type of high school attended, eighty percent of the population had less than a high school degree, but students whose parents had less than a high school degree accounted for 40% of the sample utilized in this investigation.

Table 3.12 - Comparing Student Profile to the Average Profile of the Brazilian Population

Data from 2000 Census IBGE)			Data from Socioeconomic (SES) Questionnaire and Survey from ENC	
Variable	%		Variable	%
"Skin Color" or Race *			Race	
White	53.7%		White	83.3%
Black	6.2%	44.7%	Black/Mulato	12.9%
Parda or "Mulato"	38.5%			
Yelow	0.4%		Asian	2.9%
Indigenous	0.4%		Native	0.8%
Not Informed	0.7%			
Total	100.0%		Total	100.0%
Income Intervals **			Family Income	
<= R$ 37	1.6%			
R$ 37 to R$ 75	3.9%			
R$ 75 to R$ 151	25.5%	67.7%	<= R$ 600	4.1%
R$ 151 to R$ 302	24.3%			
R$ 302 to R$ 453	12.4%			
R$ 453 to R$ 755	12.9%	24.5%	R$ 601 to R$ 2,000	28.7%
R$ 755 to R$ 1,510	11.6%			
R$ 1,510 to R$ 2,265	3.1%	5.0%	R$ 2,001 to R$ 4,000	33.2%
R$ 2,265 to R$ 3,020	1.9%			
R$ 3,020 to R$ 4,530	1.3%	2.8%	R$ 4,001 to R$ 10,000	26.5%
>= R$ 4,530	1.6%		>= R$ 10,001	7.5%
Total	100.0%		Total	100.0%
Education Level (in years) ***			Father's Education Level	
Less than 1	11.0%			
Between 1 and 3	18.9%	79.5%	Father Having less than HS Degree	39.8%
Between 4 and 7	34.3%			
Between 8 and 10	15.4%			
Between 11 and 14	15.3%		Father Having HS Degree	22.5%
At least 15	4.3%		Father Having College Degree	37.7%
Not Informed	0.9%			
Total	100.0%		Total	100.0%

* Source: IBGE 2000 Census - Table 2.1.1

** Source: IBGE 2000 Census - Table 1.7.7

*** Source: IBGE 2000 Census - Table 2.2.2

The reader, however, has to be aware that the student profile might not represent the college population or the average student who graduates from college. On one hand, the sample used in this study is a censored sample, which means that only students who were able to succeed and graduate were included in this study. On the other hand, only four fields were included in this investigation; therefore, the analysis of all students who took the ENC exam might reveal a different picture of the social inequalities of the Brazilian higher education system.

Analysis Procedure

Several types of statistical analyses were carried out in order to address the research questions posited. I used descriptive analyses (means, standard deviations, and ranges), cross-tabulation, one-way ANOVA, Kruskal Wallis test, and ordinary least-squares (OLS) regression, which are described herein. As already cited in the research rationale, data from four academic fields (Administration, Law, Civil Engineering, and Dentistry) were analyzed separately. This means that the statistical analyses described here for analyzing the data were conducted separately for each of these academic fields. By carrying out separate analyses, this investigation explored the degree to which the regression results changed among these academic fields. Due to the linear transformation of the dependent variable described early on in this chapter, not only the direction but also the magnitude of the coefficients can be compared to identify how the effects of the independent variables were associated with student test scores.

Ordinary least-squares (OLS) regressions were used to examine the relationship of some institution-level, program-level, and student-level variables to ENC student test

score. Although this investigation used data from different levels of measurements (students, programs, and institutions), OLS was selected over a hierarchical linear modeling (HLM). Osborne (2000) points out that the use of OLS regression with nested data might result in a higher probability of rejecting the null hypothesis because this statistical technique tends to produce smaller standard errors than HLM. Therefore, OLS regression will generate unbiased coefficients which provide the bases for examining the relationships addressed in this investigation. However, one should be cautious about the statistical significance of the coefficients. In order to show the differences between these regressions, Osborne ran an OLS regression and an HLM regression using a sample of students as large as 28,000 from the National Educational Longitudinal Survey of 1988. The unit of analysis (students), the magnitude of the sample size, and the dependent variable (student test score) were similar to those used in this investigation. Osborne found regression coefficients with the same direction of effect but with a significantly different statistical magnitude; the effects of SES were overestimated (4.97 in OLS versus 4.07 in HLM) and the effects of level 2 were underestimated.

The procedure applied to explore the data was the following. Multiple statistical analyses were carried out at the program level to address the first research question. Initially, correlation analyses were performed by academic field to explore the following correlations: 1) correlation among program evaluation ratings, 2) correlation between test scores and program evaluation ratings, and 3) correlation among selectivity, test scores, and program evaluation ratings. Next, mean test scores and mean program selectivity were calculated for each dimension evaluated by ACO (faculty, curriculum, and

facilities) and one-way ANOVA was used to test the existence of statistical difference between groups. In order to explore whether the relationships found in previous analyses were maintained in different institutional settings, mean test scores, mean program selectivity, and mean program evaluation ratings were determined for groups of programs defined according to the administrative control and institutional type. One-way ANOVA was used to test the existence of significant differences between groups with regard to the mean test score and mean program selectivity, while the nonparametric Kruskal Wallis test was utilized to look for significant differences in the mean program evaluation ratings.

Research question two was explored using ordinary least-squares (OLS) regression. Separate regression analysis was carried out for each academic field for the following reasons. Evaluation instruments were designed separately for each field for used in the assessment of undergraduate programs (ACO) and students (ENC). On one hand, the evaluation procedure of programs (ACO) is based on some criteria defined for each academic field. Despite the fact that the ACO outcomes represent the degree to which a program has achieved the quality standards defined by specialists of the field in the dimension assessed, the descriptive analysis and correlation analysis presented in chapter four show that the relationship between the outcomes of the ACO and ENC vary considerably across fields. On the other hand, the exit exam is designed to measure specific knowledge and skills of each academic field. As presented in Table 3.5, the means and the standard deviations differ across fields. Even after carrying out a linear transformation that made the means and standard deviations equal in all fields, I chose to

investigate the relationship within each academic field rather than in all fields together. This investigation represents the first attempt to examine the relationship between the outcomes of the ACO and the outcome of the ENC.

The results of the regression in each academic field are presented one by one in the same table to facilitate comparison of the results across fields. The variables included in this analysis were grouped into following blocks: 1) institutional characteristics, 2) evaluation of educational conditions, 3) program selectivity, 4) admissions cohort, 5) defining characteristics (age and race), 6) SES background, and 7) student involvement. Although Astin's (1991) work provides one of conceptual framework used for conducting this study, the main focus of the regression analysis was on examining the results of the full model (including all variables) rather than looking for the proportion of the variance explained by each block of variables, as Astin suggests. This relates to the fact that the variance proportion assigned to each block variables is influenced by the order they enter in the regression. The variables that are entered first in the regression are likely to have a higher proportion of the explained variance than other variables that are entered later. For instance, if a student's characteristics are entered first in the regression, followed by institutional characteristics, the influence of institutional effectiveness will be minimized because part of the explained variance will be associated with student contribution. In other words, one factor can influence the effect associated with another factor within the calculation.

In addition to the full model previously presented, a more detailed analysis was carried out in selected variables to better explain the results of the full model. Instead of

looking at the variance proportion, a series of block regression analyses were carried out to explore how initial associations with test scores changed after other variables were included in the regression equations. Unlike the first approach where the focus was to present the results of all variables in each field, this approach provided an opportunity to better understand the results of the final model presented in the final regression. The main goal was to present an in-depth analysis of how the effect of institutional characteristics, program evaluation ratings, and SES background changed as other variables were included in the analysis. These variables were selected because the findings associated with them present important implications for the policymaking process.

Conclusions

This chapter presented the research rationale that guided the investigation of the relationship between student test score and educational conditions of undergraduate programs, controlling for institutional characteristics, program selectivity, and individual characteristics. Two data sources provided most of the relevant data used in this investigation: the database containing the results of the national exit exam (ENC) and the database containing the results of the program evaluation (ACO). The dependent and independent variables were presented, along with the descriptive analyses of all data used in this investigation. Lastly, the statistical procedures used to examine the posited research questions were presented in detail.

CHAPTER IV

FINDINGS

Introduction

The results of the quantitative analysis conducted to investigate the relationship between educational conditions of undergraduate programs assessed by ACO and student performance at the ENC are presented in this chapter. First, the correlation analysis of the program-level variables is presented. Second, results are presented from the descriptive and cross-tabulation analyses designed to explore the degree to which the mean test score (ENC) and mean program selectivity were related. Third, results are presented from the descriptive and cross-tabulation analyses carried out to explore the degree to which mean test score (ENC), mean program evaluation (ACO), and mean program selectivity changed according to institutional characteristics. This analysis is important since it is very likely that not only the ACO outcomes but also mean test score and program selectivity differ considerably between programs from institutions of different type and administrative control. Fourth, the results of the regression analyses carried out within each academic field are presented in order to explore how the educational conditions of undergraduate programs, institutional type, individual characteristics, socioeconomic (SES) background were associated with student test score. Finally, this chapter closes with a summary of the most important findings. In all tables included in this chapter, the level of significance for the test carried out is represented by asterisks: three asterisks represents a p-value smaller than 0.001, two asterisks represents a p-value smaller than 0.01, and one asterisk represents a p-value smaller than 0.05.

Correlation Analysis among the Program-Level Variables

The results of the correlation analysis among the program-level variables are displayed in Table 4.1. The main interest was to explore how program evaluation ratings were related to mean test score as well as program selectivity. The correlation analyses were performed by academic field, focusing on three main relationships: 1) correlation among program evaluation ratings; 2) correlation between test scores and program evaluation ratings; and 3) correlation among selectivity, test scores, and program evaluation ratings.

Table 4.1 - Correlations among Test Scores (ENC), Program Selectivity, and Evaluation Outcomes (ACO), by Academic Fields

Academic Fields	Test Score (ENC)	Selectivity	Faculty (ACO)	Curriculum (ACO)
Administration	(N=327)			
Selectivity	0.53 ***			
Faculty (ACO)	0.33 ***	0.39 ***		
Curriculum (ACO)	0.31 ***	0.19 ***	0.44 ***	
Facility (ACO)	0.16 **	(0.03)	0.34 ***	0.51 ***
Law	(N=189)			
Selectivity	0.68 ***			
Faculty (ACO)	0.24 ***	0.32 ***		
Curriculum (ACO)	0.06	0.02	0.49 ***	
Facility (ACO)	(0.02)	(0.11)	0.22 **	0.39 ***
Civil Engineering	(N=102)			
Selectivity	0.57 ***			
Faculty (ACO)	0.64 ***	0.53 ***		
Curriculum (ACO)	0.42 ***	0.33 ***	0.46 ***	
Facility (ACO)	0.31 **	0.16	0.48 ***	0.50 ***
Dentistry	(N=80)			
Selectivity	0.54 ***			
Faculty (ACO)	0.32 **	0.39 ***		
Curriculum (ACO)	0.18	0.16	0.35 **	
Facility (ACO)	0.18	0.20	0.51 ***	0.54 ***

* p<.05 ** p<.01 *** p<.001

Correlation among Program Evaluation Scores. The results of the correlation analysis showed that program evaluation ratings were positively correlated in all fields, though with different magnitude. The highest correlations were found in Civil Engineering where program evaluation ratings were moderately correlated (correlations ranging from $r = .46$ to $r = .50$). The highest correlations in the other fields were the following: between curriculum and facility ratings in Administration ($r = .51$) between faculty and curriculum in Law ($r = .49$), and between curriculum and facility in Dentistry ($r = .54$).

Correlation between Test Scores and Program Evaluation Ratings. The results of the correlation analysis between mean test scores and program evaluation ratings are mixed, revealing important differences between fields. Although the correlation coefficients were positive in all fields but Law, only the correlations between test score and faculty rating were statistically significant in all fields; the strongest correlation was found in Civil Engineering ($r = .64$) and the weakest in Law ($r = .24$). In addition, mean test score was positively correlated (at a .05 alpha level) with all program evaluation ratings in Administration and Civil Engineering. While the correlation between test score and curriculum evaluation ratings as well as test score and facility evaluation ratings was nearly zero in Law, they were positive but nonsignificant in Dentistry.

Correlation among Selectivity, Test Scores, and Program Evaluation Ratings. As expected, program selectivity had a positive correlation with mean test scores in all fields analyzed. The strongest association was found in Law ($r = .68$), followed by Civil Engineering ($r = .57$), Dentistry ($r = .54$), and Administration ($r = .53$). In regard to the

association with program evaluation ratings, the results were mixed. Selectivity was positively and significantly correlated with faculty ratings in all fields as well as with curriculum ratings in Administration and Civil Engineering. The correlation between program selectivity and facility evaluation ratings did not reach the minimum significance level (alpha .05) in all fields. Indeed, the correlation was negative in Administration and Law.

Analyses of the Program Evaluation Ratings, Mean Test Score, and Program Selectivity

The data presented in this section provides the reader with a different view than was found in the correlation analysis, showing how the mean test scores were ordered by each evaluation rating. The number of programs, the mean test scores, and the mean program selectivity by each dimension evaluated by ACO is displayed in Table 4.2. Complete description of the data presented in this table can be found in Appendix A (more precisely in Appendix B.2.1, Appendix B.2.2, Appendix B.2.3, and Appendix B.2.4). The mean test scores and program selectivity as displayed in Table 4.2 are aggregated by each program dimension evaluated by ACO (faculty, curriculum, and facility) in order to facilitate the analysis.

Faculty Evaluation Rating. The data about faculty ratings from Table 4.2 is consistent with the correlation analysis previously presented. In all fields, the mean test score and the mean program selectivity in at least one group was significantly different from other groups, showing that programs with a high evaluation rating also had a high mean test score and high program selectivity in all fields. In addition, two important findings are noteworthy. This analysis showed a low discrimination power between

Table 4.2 - Number of Programs, Mean Test Scores (ENC), and Mean Program Selectivity, by Evaluation Outcomes (ACO) and by Academic Fields

ACO Evaluation Outcomes	Administration			Law			Civil Engineering			Dentistry		
	N	Test Score	Selec-tivity	N	Test Score	Selec-tivity	N	Test Score	Selec-tivity	N	Test Score	Selec-tivity
All programs	327	499	4.1	189	508	10.5	102	499	4.6	80	508	17.5
Faculty Evaluation		***	***		***	***		***	***		*	**
Insufficient	19	492	2.8	52	500	8.3	15	464	3.4	[a]		
Regular	166	485	3.0	63	501	9.4	40	468	2.2	17	487	11.4
Good	122	509	5.0	59	509	10.4	31	515	6.4	41	504	16.8
Very Good	20	559	9.6	15	557	22.8	16	578	8.4	22	530	23.3
Curriculum Evaluation		***	**					***	**		*	
Insufficient	56	478	3.4	60	506	11.3	[a]			12	472	14.0
Regular	111	490	3.4	47	504	8.3	28	473	3.3	24	515	17.6
Good	112	504	4.5	45	508	11.1	50	493	4.2	27	517	16.7
Very Good	48	531	5.7	37	515	11.2	24	543	7.0	17	507	20.9
Facility Evaluation		*				*		***			*	
Insufficient	28	487	4.0	47	516	13.3	[a]			11	467	15.1
Regular	75	489	4.6	26	495	8.6	26	490	4.4	34	513	15.3
Good	129	499	4.0	77	504	8.9	54	484	4.0	21	522	20.4
Very Good	95	509	4.1	39	514	11.4	22	546	6.4	14	503	20.1

[a] Due to low number of cases, programs with insufficient evaluation score were considered as with regular conditions

* p<.05 ** p<.01 *** p<.001

programs whose faculty conditions were evaluated as "insufficient" and "regular". In all cases, the difference of the mean test score and the mean program selectivity was surprisingly very close and not significantly different. Furthermore, it seems that there is a cut-off point associated with programs with "very good" faculty ratings in three fields: Administration, Law, and Civil Engineering. In these fields, programs with "very good"

evaluation ratings had a mean test score greater than 50 points in relation to programs with "good" evaluation ratings. This represents an advantage equal to one standard deviation since the standard deviation of the test scores at the program level was 50 points.

Curriculum Evaluation Rating. The results associated with mean test score and program selectivity were mixed. First, the difference between groups relative to mean test score and program selectivity was statistically significant only in Administration and Civil Engineering. In Law and Dentistry, the selectivity in programs evaluated with different levels of quality in curriculum was not significantly different. Second, a cut-off point of a half standard deviation in Civil Engineering and one-fourth of the standard deviation in Administration was found between programs whose curriculum were evaluated as "good" and "very good". Furthermore, the mean program selectivity in Law was similar in three groups while the mean test score in Dentistry did not reveal any trend.

Facility Evaluation Rating. The data from Table 4.2 illustrate in a different way the low correlation found between facility evaluation ratings and mean test score. In general, mean test score and program selectivity did not present a crescent pattern with better evaluation outcomes, a result that is consistent with the weak and nonsignificant correlations previously presented. The difference in program selectivity was not statistically significant in any field while the difference in test score was significant in all fields except Law. Indeed, programs with "insufficient" conditions in facilities had higher mean test scores and mean program selectivity than programs with better evaluation

ratings in Law. In Dentistry, programs evaluated as "regular" and "good" had higher mean test scores than programs with "very good" conditions. The exception was, once again, Civil Engineering where programs that received a "very good" evaluation rating had an advantage of the size of a half standard deviation over other programs.

Relationship between Mean Test Score (ENC), Program Selectivity, Educational Conditions (ACO), and Institutional Characteristics

This section presents the results of the analyses designed to explore whether the relationships among educational conditions and test scores change in different institutional settings. The results of the analysis related to test scores and program selectivity are presented first, followed by the results associated with educational conditions. In order to examine whether the differences in mean test scores, program selectivity, and evaluation ratings among institutional characteristics were statistically significant, one-way ANOVA and nonparametric Kruskal Wallis tests were utilized.

Mean Test Score and Program Selectivity by Institutional Characteristics. The number of programs, the mean test scores, and the mean program selectivity by institutional characteristics is presented in Table 4.3. The data reveal significant differences in test scores and program selectivity between programs from different institutional types in almost all comparisons provided. In all fields, program test scores and selectivity were higher in programs from public institutions. In regarding to administrative control, programs from federal and state institutions had the highest test scores and program selectivity, followed by non-profit institutions. In all cases, programs from for-profit institutions had the lowest means in both aspects examined. Furthermore,

Table 4.3 - Mean Test Score and Mean Program Selectivity, by Administrative Control, Institutional Type, and Academic Field

Institutional Characteristics	Administration			Law			Civil Engineering			Dentistry		
	N	Test Score	Selec-tivity	N	Test Score	Selec-tivity	N	Test Score	Selec-tivity	N	Test Score	Selec-tivity
All Programs		498.8	4.1	189	507.7	10.5	102	499.1	4.6	80	507.5	17.5
Public x Private		***	***		***	***		***	***		***	***
Private	245	490.6	2.6	135	489.8	6.7	52	466.7	2.0	40	485.3	10.8
Public	82	523.2	8.7	54	552.4	19.8	50	532.7	7.4	40	529.8	24.2
Administrative Control		***	***		***	***		***	***		***	***
Federal	36	542.7	9.9	35	563.5	22.0	32	541.6	7.1	27	534.0	21.1
State	30	517.8	10.0	11	548.1	19.9	15	528.8	8.9	12	525.6	32.2
Local	16	489.3	3.4	8	510.0	10.0	3	457.5	1.9	1	466.0	9.9
Relig/Community	115	504.5	2.5	76	493.8	6.4	33	473.9	2.1	24	486.3	11.7
Private for Profit	130	478.3	2.7	59	484.7	7.2	19	454.4	1.9	16	483.6	9.3
Institutional Type		***	***		**	**			*			
University	164	513.5	5.2	121	516.4	12.0	82	505.8	5.2	70	510.2	18.1
Univ. Center	40	484.4	2.9	16	484.6	7.4	8	461.0	1.9	3	456.6	14.5
College	123	483.8	3.1	52	494.6	7.9	12	479.0	2.8	7	502.5	12.7

Note: One-Way ANOVA was used to compare means
* p<.05 ** p<.01 *** p<.001

programs from universities had significantly higher test scores and selectivity, followed by colleges and university-centers with very similar mean test scores in all fields but Dentistry. However, no significant differences (at a p-value .01) were found between test scores and program selectivity of those programs from universities, university-centers, and colleges in Civil Engineering and Dentistry. This finding is consistent with the fact that the majority of programs in these fields are from universities, as already discussed.

Program Evaluation Ratings by Institutional Characteristics. The data provided in Table 4.4 address the relationship between the educational conditions assessed by

Table 4.4 - Mean ACO Outcome for Faculty, Curriculum, and Facility, by Administrative Control, Institutional Type, and Academic Field

Institutional Characteristics	Administration			Law			Civil Engineering			Dentistry		
	Fac.	Cur.	Inf.	Fac.	Cur.	Inf.	Fac.	Cur.	Inf.	Fac.	Cur.	Inf.
All Programs	2.44	2.46	2.89	2.20	2.31	2.57	2.47	2.94	2.90	3.04	2.61	2.48
Public x Private	***		***	***			***	*		***		
Private	2.35	2.45	3.01	2.02	2.25	2.66	2.04	2.79	2.96	2.75	2.65	2.50
Public	2.70	2.50	2.52	2.63	2.46	2.35	2.92	3.10	2.84	3.33	2.58	2.45
Administrative Control	***	**	***	***			***	**		*		
Federal	3.03	2.83	2.56	2.83	2.60	2.31	3.00	3.28	2.91	3.30	2.52	2.33
State	2.60	2.40	2.47	2.64	2.27	2.64	3.00	3.00	2.80	3.42	2.83	2.75
Local	2.13	1.94	2.56	1.75	2.13	2.13	1.67	1.67	2.33	3.00	1.00	2.00
Relig/Community	2.43	2.63	3.21	2.08	2.38	2.74	2.03	2.82	3.03	2.79	2.67	2.54
Private for Profit	2.28	2.30	2.84	1.95	2.08	2.56	2.05	2.74	2.84	2.69	2.63	2.44
Institutional Type	***	***	***	***	***		*					
University	2.65	2.68	3.02	2.45	2.54	2.64	2.60	3.00	2.93	3.09	2.67	2.53
Univ. Center	2.58	2.55	3.25	2.38	2.19	2.81	2.13	2.63	2.88	3.33	2.00	2.67
College	2.11	2.15	2.59	1.56	1.83	2.35	1.83	2.75	2.75	2.43	2.29	1.86

Note: Kruskal Wallis nonparametric test was used to compare the results
* $p<.05$; ** $p < 0.01$; *** $p < 0.001$

ACO (faculty, curriculum, and facility) and institutional characteristics. In regard to the differences between programs from public and private institutions, programs from public institutions received higher evaluation ratings in faculty and curriculum in all fields analyzed than did programs from private institutions. In contrast, programs from private institutions had higher facility evaluation ratings in all fields and higher curriculum evaluation ratings in Dentistry.

Regarding the administrative control of institutions that maintained undergraduate programs analyzed, the results were somewhat mixed. Programs from federal institutions

showed the highest faculty and curriculum ratings in Administration, Law, and engineering while programs from state institutions received the highest ratings in all dimensions assessed in Dentistry. Another important finding relates to the results of the educational facility evaluation. Programs in non-profit institutions received the highest faculty evaluation ratings in all fields but Dentistry, although only in Administration were differences between groups significant.

The results also showed that programs from universities had superior educational conditions. Universities programs were evaluated higher than others in faculty and curriculum in all fields except Dentistry, although the evaluation ratings were significantly different only in Administration and Law. On the contrary, colleges had the lowest means in all three dimensions assessed in all fields. The exception was curriculum evaluation ratings in Civil Engineering and Dentistry. It is important to point out that the insignificant difference found in Civil Engineering and Dentistry is likely to be due to the strong concentration of programs in universities, which resulted in a low number of cases in smaller institutions, especially in Dentistry.

Regression Analyses of the Student-Level Variables

This section is divided into two parts. The first part presents the results of the regression carried out separately for each academic field. The main focus is to explore the effect of each block of independent variables on student performance. The second part presents the results of the blocked regression carried out to explore results of selected variables on student performance. Due to the linear transformation carried out to standardize student test scores, the coefficients are to be interpreted in relation to the size

of the standard deviation of the distribution, which is 100.0 in all academic fields due to the linear. Therefore, a coefficient of -50.0, for instance, means that the effect of a given variable was equivalent to a half standard deviation behind the reference group.

Exploring the Effect of Each Block of Independent Variables on Student Performance

In this section, the effect associated with institutional characteristics, educational conditions of programs, student's characteristics, student involvement, and program selectivity upon test score is presented. The coefficients, the level of significance, and the standard errors of all variables from the final regression model are displayed in Table 4.5. The result of the regression by each block of variables is presented in the following order: 1) institutional characteristics, 2) evaluation of educational conditions, 3) program selectivity, 4) admission cohort, 5) SES background, 6) defining characteristics (age and race), and 7) student involvement.

The Effect of Institutional Characteristics on Student Performance. The regression revealed important differences between the influence of institutional type and administrative control on test scores. Regarding the type of administrative control, the effect associated with a program from a federal or state institution was positive and greater than other programs from state, non-profit, or for-profit institutions in Law, Civil Engineering, and Dentistry. In Administration, however, the strongest effect on test scores was found in programs from non-profit institutions (religious or community oriented institutions). The negative coefficient associated with being a program from a

Table 4.5 - Regression Coefficients from Full Model, by Academic Field

Variables	Administration		Law		Civil Engineering		Dentistry	
	Beta	SE	Beta	SE	Beta	SE	Beta	SE
	(N=30776)		(N=34834)		(N=4880)		(N=6595)	
Institutional Characteristics								
Administrative Control								
Federal Institution	12.83 ***	2.89	52.69 ***	2.57	44.22 ***	5.75	37.29 ***	4.69
State Institution	-22.82 ***	3.12	31.37 ***	3.32	32.43 ***	6.71	19.30 **	6.34
Local Institution	12.47 ***	2.71	27.32 ***	2.71	17.81 *	8.23	-1.67	11.88
Religious/Community	24.89 ***	1.37	20.19 ***	1.27	7.83	4.71	14.33 ***	3.21
(Reference Group=For-Profit Institutions)								
Institutional Type								
University	-5.10 ***	1.49	-11.98 ***	1.50	13.21 *	5.37	0.41	4.56
University Center	-6.21 ***	1.69	-8.93 ***	2.02	17.50 *	7.04	-21.83 **	7.79
(Reference Group=Colleges)								
Program Characteristics								
Evaluation of Educational Conditions								
Faculty								
Insufficient	-1.25	4.49	-5.68 *	2.48	-61.80 ***	7.06		
Regular	-17.61 ***	2.50	-11.88 ***	2.25	-40.44 ***	6.16	-15.45 ***	4.64
Good	-15.45 ***	2.31	-13.67 ***	2.05	-31.58 ***	4.54	-15.94 ***	3.31
(Reference Group=Very Good Evaluation Outcome)								
Curriculum								
Insufficient	-20.45 ***	2.49	-3.34	1.91			2.36	4.78
Regular	-12.15 ***	1.77	0.45	1.79	-29.66 ***	4.90	14.75 ***	3.85
Good	-11.07 ***	1.70	-1.97	1.58	-14.23 ***	3.70	13.66 ***	3.66
(Reference Group=Very Good Evaluation Outcome)								
Facility								
Insufficient	8.98 **	3.27	5.16 **	1.85			-0.30	5.69
Regular	-1.77	1.93	-11.40 ***	2.10	2.79	5.30	24.38 ***	4.44
Good	2.19	1.30	3.31 *	1.40	-5.25	3.85	16.51 ***	3.88
(Reference Group=Very Good Evaluation Outcome)								
Program Selectivity								
Selection Ratio	8.28 ***	0.22	3.05 ***	0.11	1.28 *	0.57	1.63 ***	0.19
Region								
Central	36.25 ***	3.96	24.73 ***	3.59	2.00	9.09	64.59 ***	11.90
Northeast	20.67 ***	3.85	25.29 ***	3.28	23.62 **	7.79	44.96 ***	10.41
Southeast	47.47 ***	3.54	16.16 ***	3.08	26.18 **	8.07	38.36 ***	10.62
South	57.14 ***	3.76	27.32 ***	3.32	21.33 *	8.35	72.47 ***	10.84
(Reference Group=North)								

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 4.5 - Regression Coefficients from Full Model, by Academic Field

Variables	Administration		Law		Civil Engineering		Dentistry	
	Beta	SE	Beta	SE	Beta	SE	Beta	SE
Admission Cohort								
1995 or Before	-7.43 ***	1.93	-35.15 ***	1.92	-35.17 ***	3.62	-51.06 ***	7.37
1996	-6.79 ***	1.82	-10.76 ***	1.25	-25.24 ***	3.06	-14.27 ***	3.93
1998 or After	-3.55 **	1.29	-15.66 ***	3.29	-7.80	8.15	-5.95 *	2.75
(Reference Group=1997 Cohort)								
Individual Characteristics								
Defining Characteristics								
Age Group								
25-30	-1.99	1.29	-21.32 ***	1.28	-24.98 ***	3.18	-26.38 ***	3.28
>30	14.44 ***	1.55	-17.64 ***	1.42	-35.17 ***	5.25	-8.90	8.69
(Reference Group=Younger than 25)								
Race								
Black / Mulato	-9.82 ***	1.62	-1.57	1.52	-18.06 ***	3.61	1.67	4.04
Asian	0.55	2.92	6.75 *	3.44	12.50	7.05	-15.75 **	5.08
Native	-9.83	6.19	-14.00 **	5.34	-30.72 *	12.96	-14.75	17.62
(Reference Group=White)								
SES Background								
Income								
R$ 601 to R$ 2.000	14.94 ***	2.79	9.01 ***	2.52	-1.18	5.96	5.22	7.86
R$ 2.001 to R$ 4.000	26.59 ***	2.82	16.62 ***	2.52	3.78	6.04	5.00	7.74
R$ 4.001 to R$ 10.000	41.58 ***	2.95	19.38 ***	2.58	-0.11	6.28	5.56	7.84
>= R$ 10.001	43.08 ***	3.47	16.55 ***	3.04	12.61	7.45	-5.31	8.64
(Reference Group=Income less than R$ 601)								
Father's Education								
HS	0.76	1.39	-1.17	1.40	1.72	3.41	3.44	3.42
College	5.64 ***	1.49	1.49	1.32	5.98	3.38	-1.22	3.08
Type of HS Attended								
All Pub.	1.37	1.31	3.68 **	1.25	11.26 ***	3.05	8.76 *	3.43
Public/Private	-15.89 ***	1.63	-13.32 ***	1.46	-10.65 **	3.88	-6.04	3.44
Student Involvement	5.37 ***	0.39	7.49 ***	0.38	3.27 **	1.00	9.75 ***	1.06
(Constant)	396.41		427.50		497.05		357.67	
Adjusted R Square	0.16		0.17		0.30		0.18	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

state institution requires further explanation that involves the effect of program selectivity, which is addressed later.

In regard to the effects associated with institutional type, the results were mixed. In Civil Engineering, being a program from a university or university-center was a significant, positive predictor of student test scores. In the other fields, however, the results were the opposite. Surprisingly, the effect associated with programs from a university or university-center was negative in other fields, suggesting that small higher education institutions (such as colleges) provide an environment associated with higher student performance. However, it is important to consider that there are other aspects such as program selectivity and administrative control that play a very important role in the interpretation of results.

The Effect of Program Evaluation Ratings on Student Performance. The regression results of the educational conditions assessed by ACO upon test score showed very interesting and unexpected findings. The evaluation rating of "very good" was selected as the reference group in order to compare other programs evaluated with "good", "regular", and "insufficient" conditions. Therefore, coefficients would be expected to be negative if the better educational condition was associated with higher test scores in all categories.

The results of the faculty evaluation ratings showed the most consistent and strongest results among the educational conditions evaluated by ACO. Overall, the higher the evaluation rating, the higher the test score. As can be seen in Table 4.5, all coefficients associated with faculty evaluations were negative and statistically significant

in all fields except Administration. This finding indicates that the effect of programs with "insufficient", "regular", and "good" evaluation ratings on student test scores was negative compared to programs evaluated with "very good" faculty ratings. The strongest effect was found in engineering, followed by Administration and Dentistry. In Civil Engineering, the negative effect of "insufficient" faculty condition was equivalent to a half standard deviation behind programs very well evaluated (beta = -61.80).

The effect of curriculum evaluation ratings on student test scores was mixed. Curriculum evaluation rating was a significant predictor of student performance in Administration, Civil Engineering, and Dentistry. In Administration and Civil Engineering, a better curriculum evaluation rating corresponded directly to a higher test score. Surprisingly, the effect associated with quality of curriculum in Dentistry had opposite results---the effect of programs with "regular" and "good" evaluation ratings was greater than the effect of programs with "very good" evaluation rating (the reference group). In regard to Law, curriculum rating was not a significant predictor of test score in Law, as the size of coefficients was non-significant and nearly zero.

Interestingly, the educational conditions of facilities presented very confusing results. In Administration and Civil Engineering, the educational condition of facilities was an insignificant predictor of test scores after all variables were controlled. In other fields, the results were either contrary to expectations (e.g., in Dentistry) or inconsistent (e.g., in Law). For instance, the effect of programs with "regular" and "good" evaluation ratings in Dentistry and "insufficient" and "good" evaluation ratings in Law was greater than the effect of programs with "very good" evaluation ratings.

The Effect of Program Selectivity on Student Performance. Selectivity was a significant predictor of student test scores in all fields, although the magnitude of its effect changed between fields. The biggest effect of selectivity was found in Administration, followed by Law, Dentistry, and Civil Engineering. Despite the relative small size of the coefficient associated with program selectivity in Civil Engineering and Dentistry, it is important to remember that it relates to one unit change in the selection ratio. Therefore, when the selection ratio is considered, the total effect of selectivity tends to be considerable, especially in programs from federal and state universities, as displayed in Table 4.3. In order to mention the magnitude of the effect associated with selectivity, it is necessary to consider the difference between the mean selectivity of a program from a federal or a state institution and the mean selectivity of a program from a for-profit institution (the reference group in the regression analysis) and then multiply by the coefficient in each field. For instance, the advantage of a program from a federal institution considering only the difference in the selectivity was 59.6 , 45.1, 6.7, and 19.2 in Administration, Law, Civil Engineering, and Dentistry, respectively. In regard to the results associated with state institutions, the results were in the same magnitude except in Civil Engineering where the advantage was even bigger (37.3). In addition, special attention is placed on the degree to which the inclusion of this selectivity variable affects the size of coefficients of the other variables, such as institutional characteristics, program evaluation ratings, and SES background.

The Effect of Length of Time in College. The results associated with the admission cohort were very interesting. Considering the fact that all students took the ENC exam in

2001 (the year they were expected to graduate), the admission cohort can be used to determine the approximate time (in years) to graduation. Therefore, the effect associated with graduating in 4 years (1998 cohort), 6 years (1996 cohort), and 7 years (1995 cohort) was negative compared to students graduating in 5 years (1997 cohort) in all fields analyzed. This finding is significant from the policymaking standpoint, revealing a negative association with the number of years students take to achieve graduation and their performance on this test. As can be seen in Table 4.5, the magnitude of the coefficients associated with the 1995 admission cohort (7 years to graduate) in Law, Civil Engineering, and Dentistry was considerably bigger than the other groups. It is also noteworthy to discuss the effect associated with a 4-year program versus a 5-year program in Administration and Dentistry. The results suggest that programs with 5-year periods offer stronger preparation than programs with 4-year length, possibly due to the additional number of courses that they are able to offer as a result of having an additional year.

The Effect of SES Background on Student Performance. The effect of student's SES background on test scores was somehow surprising. Among the SES background variables included in the regression, it seems that parental income had the strongest effect on test scores only in two fields: Administration and Law. In Civil Engineering and Dentistry, the nonsignificance of the coefficients suggests that parental income did not affect the test score of those students who persisted in college and took the ENC exam. However, the size of the coefficient in Civil Engineering indicates that the students from upper income level scored higher than other students while the sign of the coefficient in

dentistry indicates that students from upper income level scored lower than other students.

In regard to other measures of SES background, the results were even more surprising. For instance, father's education was not a significant predictor of test score across all fields except Administration. Comparing to students who attended public high school, the effect was negative in all fields and significant in all fields except Administration. This finding contradicts the common myth that private high schools better prepare students for college. Furthermore, students who moved from public to private high schools (and vice-versa) had inferior performance in the ENC test. These findings required further investigation that is provided in the following section of this chapter.

The Effect of Other Individual Characteristics on Student Performance. In order to control for some individual characteristics, two groups of variables were included in the regression. The first group of dummy variables refers to age group. The results indicated that older students achieved lower test scores than younger students in Law, Civil Engineering, and Dentistry, which might be related to two aspects. The first aspect is that the admission to college happened immediately after graduating from HS, which might be an indication of best students being admitted to college directly from high school. The second aspect is related to the time taken to graduate from college, which may be a sign that students who graduate on time leave college better prepared. In Administration, it was surprising to find that the effect associated with being older than

30 years of age was positive, while no difference was found between students from the other two age groups.

In regard to the effect of race on test score, being a member of underrepresented racial groups (Black/Mulatto as well as Native) represented a negative effect on test score in all fields but Dentistry. In Dentistry, the effect associated with being a Black/Mulatto student was not significantly different from Whites. This finding might be related to the selectivity of the field as well as the low percentage of Black students in this field (only nine percent). In regard to students with Asian background, the results of the regressions showed that they had a significantly negative effect on test score in Dentistry, a positive but nonsignificant effect in Law and Civil Engineering, and no difference in Administration.

The Effect of Student Involvement in College on Student Performance. The results of the analysis of student involvement in certain activities in college showed interesting results. Test scores were higher for those students who responded to be involved in some activities (e.g., going to the library, looking for out-classroom orientation, and participating in extra academic activities, extra-curricular activities, and events). The results were consistent in all academic fields analyzed, having the strongest effect in Dentistry. Despite the small magnitude of the coefficients, this finding suggests the importance of student involvement in certain college activities as a means to help them develop and learn. Although the main focus of this investigation was to examine institutional performance, student involvement was also included in the analysis to explore the degree to which it influenced student performance at the ENC. In order to

verify the influence of the inclusion of this variable in the regression, two regressions were carried out (one with all variables except student involvement and another with all variables) and their results were compared. It is important to point out that the inclusion of this variable did not affect the results of other variables.

Exploring the Results of the Regression Analysis of Selected Variables

This section provides the reader with a better understanding about the results of institutional characteristics, educational conditions, and type of high school attended. The effects of these variables is examined in a more detailed manner in this section because they have significant implications for policy making (specially related to the stratification of the Brazilian higher education system) as discussed in the last chapter. A blocked regression was carried out for each of these variables to show how the initial advantages (or disadvantages) changed as other variables were incorporated into the analysis.

The Effect of Administrative Control on Student Performance

One of the goals of this investigation was to explore the influence of different institutional settings on student performance. In order to better understand the results associated with administrative control, a series of regressions was carried out to show how the effect associated with administrative control changed after other variables were included in the analysis. The variables were collapsed into six blocks containing the following information: 1) administrative control (federal, state, local, religious/community institutions), 2) institutional type (university, university-center, and colleges), 3) student's characteristics (admissions cohort, age, race, income, type of high

school, and father's education), 4) evaluation of educational conditions (faculty, curriculum, and facility evaluation ratings), 5) student involvement, and 6) program selectivity. Program selectivity was included in the last block because its inclusion in the regression caused a lot of changes in other coefficients, which is discussed later. The results of the blocked regression displayed in Table 4.6 revealed four important findings.

First, the initial advantage of programs from federal and state institutions (as can be seen in Model A) was impressive in all fields examined. The advantage of federal institutions over private for-profit institutions was greater than two-thirds of the standard deviation in all fields examined. In regard to state institutions, its advantage varied between one-third of the standard deviation in Administration to almost four-fifths in dentistry and Civil Engineering. In Administration, programs from state institutions had similar average test scores compared to programs from private non-profit institutions. It is also important to point out that non-profit institutions (religious-oriented and community oriented) had higher test scores than local institutions and for-profit institutions in all fields except Law.

Second, the initial advantage of programs from federal and state institutions remained large even after controlling for other variables except program selectivity. Overall, no substantial changes in the coefficients were found when student performance was controlled for institutional type (Model B), student's characteristics (Model C), and student involvement (Model E). However, when the quality of programs was included in the analysis (Model D), part of the advantage of state institutions in Civil Engineering as well as federal institutions in Administration and Civil Engineering was substantially

Table 4.6 - Regression Coefficients for Administrative Control

Administrative Control	Model A Adm. Control Only	Model B Adding Instituio-nal Type	Model C Adding Student Charate-ristics	Model D Adding Program Evaluation	Model E Adding Student Involve-ment	Model F Adding Program Selectivity
Administration						
Federal Institution	70.71 ***	73.00 ***	77.81 ***	62.50 ***	60.81 ***	12.83 ***
State Institution	31.37 ***	32.96 ***	40.14 ***	40.34 ***	38.49 ***	-22.82 ***
Local Institution	10.25 ***	9.80 ***	15.66 ***	20.76 ***	20.51 ***	12.47 ***
Religious/Comm.	29.48 ***	30.98 ***	31.61 ***	30.59 ***	29.91 ***	24.89 ***
(Reference Group=For-Profit Institutions)						
Law						
Federal Institution	94.42 ***	101.18 ***	101.34 ***	98.25 ***	96.00 ***	52.69 ***
State Institution	78.28 ***	84.27 ***	81.24 ***	82.23 ***	80.19 ***	31.37 ***
Local Institution	35.00 ***	33.61 ***	30.01 ***	30.43 ***	30.46 ***	27.32 ***
Religious/Comm.	19.87 ***	23.80 ***	24.53 ***	24.22 ***	23.82 ***	20.19 ***
(Reference Group=For-Profit Institutions)						
Civil Engineering						
Federal Institution	77.87 ***	77.65 ***	72.40 ***	48.51 ***	46.55 ***	44.22 ***
State Institution	79.37 ***	78.85 ***	66.39 ***	45.31 ***	42.80 ***	32.43 ***
Local Institution	-1.34	3.81	2.81	21.90 **	20.19 *	17.81 *
Religious/Comm.	2.37	7.98	7.12	9.78 *	9.05 *	7.83
(Reference Group=For-Profit Institutions)						
Dentistry						
Federal Institution	64.03 ***	64.67 ***	67.71 ***	62.33 ***	58.91 ***	37.29 ***
State Institution	55.44 ***	54.57 ***	58.07 ***	59.36 ***	55.85 ***	19.30 **
Local Institution	-6.97	-7.84	-9.58	0.80	-1.56	-1.67
Religious/Comm.	9.04 **	8.57 **	10.24 ***	17.02 ***	16.86 ***	14.33 ***
(Reference Group=For-Profit Institutions)						

* $p<0.05$; ** $p < 0.01$; *** $p < 0.001$

Model A: Only administrative control (federal, state, local, religious/community institutions)

Model B: Adding institutional type (university, university-center, and colleges)

Model C: Adding admission cohort, age, race, income, type of high school, and father's education

Model D: Adding program evaluation (faculty, curriculum, and facility evaluation ratings)

Model E: Adding student involvement

Model F: Adding program selectivity

reduced due to the superior quality of their programs. For instance, in Civil Engineering, the coefficient size of programs from federal institutions and state institutions was reduced by thirty-two percent when quality of programs was included in the regression.

Third, the results showed that a substantial part of the advantage of federal and state institutions might be due to the selectivity of their programs. The advantage of programs from federal institutions was substantially reduced in Administration (79%), Law (45%), and Dentistry (37%) while remaining in the same magnitude as Civil Engineering when selectivity was added to the analysis. The advantage associated with programs from state institutions showed substantially more changes than programs from federal institutions. The advantage of programs from these institutions in Law and Dentistry was reduced to one-third and a substantial disadvantage was revealed in Administration when selectivity was included.

Lastly, the effect associated with private non-profit institutions (religious and community institutions) did not substantially change after other variables were included in the analysis. When administrative control was regressed alone on student performance (Model A), students from these institutions presented inferior test scores compared to their counterparts from federal and state institutions and superior test scores compared to for-profit institutions. Interestingly, these institutions (community and religious oriented institutions) had the strongest positive effect on test score in Administration when program selectivity was added to the regression (Model F). However, this result can be explained by the fact that the coefficients associated with federal and state institutions dramatically dropped when selectivity was controlled.

The Effect of Institutional Type on Student Performance

Similar to the analysis carried out to explore the effect of types of administrative control on student performance, this section presents the results of the block regression conducted to demostrate the effect associated with institutions of different types (e.g., university, university-center, and colleges) and associated changes after other variables were included in the analyses. The main focus of this analysis is to verify how the initial difference among programs from universities, university-centers, and colleges was transformed in the final results presented in the previous section. In order to achieve that, six blocks of variables were created containing the following information: 1) institutional type (university, university-center, and colleges), 2) administrative control (federal, state, local, religious/community institutions), 3) student's characteristics (admissions cohort, age, race, income, type of high school, and father's education), 4) evaluation of educational conditions (faculty, curriculum, and facility evaluation ratings), 5) student involvement, and 6) program selectivity. The results of this regression displayed in Table 4.7 revealed two important findings.

First, programs from universities had a higher mean test score than programs from university-centers and colleges in all fields examined (Model A). It is important to note, however, that the advantage of universities was small in three fields examined: Administration, Law, and Dentistry. In these fields, the initial positive effect of programs from universities turned into a negative effect after administrative control was included in the analysis (Model B). Even in Civil Engineering where universities had a considerably high initial advantage equivalent to two-thirds of the standard deviation over programs

Table 4.7 - Regression Coefficients for Institutional Type

Institutional Type	Model A Institutio-nal Type Only	Model B Adding Adm. Control	Model C Adding Student Characte-ristics	Model D Adding Program Evaluation	Model E Adding Student Involve-ment	Model F Adding Program Selectivity
Administration						
University	14.22 ***	-4.46 **	-9.65 ***	-14.34 ***	-13.82 ***	-5.10 ***
University Center	-5.00 **	-1.93	-5.22 **	-9.04 ***	-8.77 ***	-6.21 ***
(Reference Group=Colleges)						
Law						
University	8.26 ***	-13.12 ***	-13.76 ***	-14.68 ***	-13.52 ***	-11.98 ***
University Center	-6.08 **	-2.72	-5.69 **	-3.77	-3.09	-8.93 ***
(Reference Group=Colleges)						
Civil Engineering						
University	66.85 ***	25.90 ***	22.78 ***	7.69	7.90	13.21 *
University Center	27.16 ***	34.83 ***	25.11 ***	16.92 *	16.40 *	17.50 *
(Reference Group=Colleges)						
Dentistry						
University	11.85 **	-5.30	-4.89	-3.41	-2.50	0.41
University Center	-23.07 **	-39.26 ***	-39.21 ***	-25.18 **	-24.61 **	-21.83 **
(Reference Group=Colleges)						

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Model A: Only institutional type (university, university-center, and colleges)

Model B: Adding administrative control (federal, state, local, religious/community institutions)

Model C: Adding admission cohort, age, race, income, type of high school, and father's education

Model D: Adding program evaluation (faculty, curriculum, and facility evaluation ratings)

Model E: Adding student involvement

Model F: Adding program selectivity

from colleges, the inclusion of administrative control in the regression caused a reduction of 66% in its coefficient (see change from Model A to Model B).

Second, programs from university-centers had an inferior mean test score than programs from colleges (Model A) in all fields except Civil Engineering. Unlike the effect associated with universities, the inclusion of other variables in the analysis did not significantly change the coefficients in all fields. This means that the initial disadvantage associated with programs from university-centers was maintained even after other variables were included in the analysis.

The Effect of Educational Conditions on Student Performance

The major goal of this investigation was to explore the influence of the educational conditions assessed by the ACO on student performance. Following the same strategy presented in the previous section, a series of regressions was carried out to show how the effect associated with each dimension of the educational conditions (e.g., faculty, curriculum, and facilities) changed after other variables were included in the analysis. Six blocks of variables were created containing the following information: 1) the dimension of the educational condition being explored (e.g., faculty evaluation ratings), 2) the other two dimensions of the educational conditions (e.g., curriculum and facility evaluation ratings), 3) administrative control (federal, state, local, religious/community institutions) and institutional type (university, university-center, and colleges), 4) student's characteristics (admissions cohort, age, race, income, type of high school, and father's education), 5) student involvement, and 6) program selectivity. The results associated

with faculty are presented first, followed by the results associated with curriculum and facilities.

Faculty Educational Conditions. The blocked regression was carried out to explore the effect of faculty on student performance and results are displayed in Table 4.8. The difference in mean test scores from programs with low evaluation ratings were considerably high in all fields examined (Model A). As other variables were included in the analysis, the initial advantage associated with programs with "very good" faculty evaluation ratings was substantially reduced. However, the coefficients remained negative throughout all models that controlled for institutional characteristics, student's characteristics, student involvement, and program selectivity. Particular attention has to be placed on the degree to which the inclusion of institutional characteristics, student's characteristics, and program selectivity affected the size of the coefficients of faculty educational conditions.

The inclusion of institutional characteristics (e.g., institutional type and type of administrative control) on the regression equation caused a reduction in the size of the coefficients by approximately 20% in Administration, 35% in Civil Engineering, and more than 50% in Law and Dentistry. The changes observed are consistent wiuth the fact that the faculty educational conditions vary significantly across types of institutions, as can be seen in Table 4.4. Therefore, the disadvantage associated with lower faculty ratings was considerably reduced in all fields examined when institutional characteristics (especially type of administrative control) were added to the regression. This means that part of the disadvantage associated with programs evaluated with "insufficient"

Table 4.8 - Regression Coefficients for Faculty Educational Conditions

Faculty Educational Conditions	Model A Faculty Evaluation Ratings	Model B Adding Curric. And Facility Evaluation Ratings	Model C Adding Inst. Type and Adm. Control	Model D Adding Student Characte-ristics	Model E Adding Student Involve-ment	Model F Adding Program Selectivity
Administration						
Insufficient	-69.28 ***	-45.58 ***	-34.95 ***	-21.83 ***	-23.43 ***	-1.25
Regular	-74.53 ***	-64.72 ***	-53.01 ***	-43.81 ***	-43.06 ***	-17.61 ***
Good	-57.46 ***	-50.64 ***	-43.21 ***	-37.31 ***	-36.49 ***	-15.45 ***
(Reference Group=Very Good Evaluation Outcome)						
Law						
Insufficient	-49.20 ***	-45.51 ***	-14.36 ***	-6.95 **	-7.74 **	-5.68 *
Regular	-47.68 ***	-48.98 ***	-19.39 ***	-12.55 ***	-13.99 ***	-11.88 ***
Good	-51.70 ***	-47.79 ***	-22.86 ***	-16.58 ***	-17.40 ***	-13.67 ***
(Reference Group=Very Good Evaluation Outcome)						
Civil Engineering						
Insufficient	-127.91 ***	-119.94 ***	-77.60 ***	-66.15 ***	-66.38 ***	-61.80 ***
Regular	-113.50 ***	-99.52 ***	-60.29 ***	-46.83 ***	-47.30 ***	-40.44 ***
Good	-63.76 ***	-60.38 ***	-40.93 ***	-37.30 ***	-37.64 ***	-31.58 ***
(Reference Group=Very Good Evaluation Outcome)						
Dentistry						
Regular	-38.09 ***	-45.16 ***	-13.05 **	-13.80 **	-16.11 ***	-15.45 ***
Good	-27.11 ***	-32.93 ***	-14.83 ***	-16.25 ***	-17.69 ***	-15.94 ***
(Reference Group=Very Good Evaluation Outcome)						

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Model A: Only faculty evaluation ratings

Model B: Adding curriculum, and facility evaluation ratings

Model C: Adding institutional type (university, university-center, and colleges) and administrative control (federal, state, local, religious/community institutions)

Model D: Adding admission cohort, age, race, income, type of high school, and father's education

Model E: Adding student involvement

Model F: Adding program selectivity

conditions is due to certain characteristics associated with the type of administrative control. Yet, consistency of the effect across academic fields suggests that faculty seem to be a critical resource for student success in college.

Another important finding relates to the degree to which the coefficients changed after student's characteristics were added to the regression equation. The change in the coefficients between models D and C revealed that the negative effect associated with faculty ratings was considerably reduced in Administration and Law. In these fields, the socioeconomic (SES) background, especially family income, had a strong effect on test scores, as revealed in Table 4.5.

Care must be taken when analyzing the results of the final model (F) in the Administration field. The inclusion of program selectivity in the regression analysis drastically reduced the coefficients associated with faculty evaluation ratings. This substantial reduction of coefficients in Administration might indicate that the candidates-to-spaces ratio used to control for student quality might not be an adequate measure.

Curriculum Educational Conditions. The results of the regression displayed in Table 4.9 provide the basis to better understand the mixed results associated with curriculum evaluation ratings. The unexpected effect of curriculum on test scores in Law and Dentistry reflects the initial difference between curriculum evaluation ratings in these fields. Even when only curriculum was included in the analysis (Model A), very little difference was found in Law, while programs with "regular" and "good" evaluation ratings had mean test scores superior to programs with "very good" evaluation ratings in Dentistry. In the other fields, the initial effect of curriculum was reduced by about two-

Table 4.9 - Regression Coefficients for Curriculum Educational Conditions

Curriculum Educational Conditions	Model A Curriculum Evaluation Ratings	Model B Adding Faculty and Facility Evaluation Ratings	Model C Adding Inst. Type and Adm. Control	Model D Adding Student Characte-ristics	Model E Adding Student Involve-ment	Model F Adding Program Selectivity
Administration						
Insufficient	-55.76 ***	-43.52 ***	-33.41 ***	-32.80 ***	-32.04 ***	-20.45 ***
Regular	-35.45 ***	-23.56 ***	-15.90 ***	-14.48 ***	-14.20 ***	-12.15 ***
Good	-38.57 ***	-29.15 ***	-19.60 ***	-18.17 ***	-17.49 ***	-11.07 ***
(Reference Group=Very Good Evaluation Outcome)						
Law						
Insufficient	-4.88 **	-0.12	-4.00 *	-5.59 **	-5.68 **	-3.34
Regular	1.03	10.24 ***	1.59	-0.57	-0.58	0.45
Good	-2.92	0.40	3.55 *	1.35	2.12	-1.97
(Reference Group=Very Good Evaluation Outcome)						
Civil Engineering						
Regular	-93.05 ***	-36.47 ***	-31.68 ***	-29.97 ***	-28.38 ***	-29.66 ***
Good	-47.90 ***	-15.03 ***	-9.90 **	-11.80 ***	-11.18 **	-14.23 ***
(Reference Group=Very Good Evaluation Outcome)						
Dentistry						
Insufficient	-24.80 ***	-9.52 *	-3.65	-3.91	-3.95	2.36
Regular	11.03 **	11.80 **	5.38	6.50	6.11	14.75 ***
Good	16.90 ***	15.73 ***	13.72 ***	14.26 ***	12.80 ***	13.66 ***
(Reference Group=Very Good Evaluation Outcome)						

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Model A: Only curriculum evaluation ratings

Model B: Adding faculty and facility evaluation ratings

Model C: Adding institutional type (university, university-center, and colleges) and administrative control (federal, state, local, religious/community institutions)

Model D: Adding admission cohort, age, race, income, type of high school, and father's education

Model E: Adding student involvement

Model F: Adding program selectivity

thirds as other variables were included in the analysis. Additionally, the coefficient sizes were significantly reduced in Administration and Civil Engineering when faculty ratings and facility ratings (Model B) as well as institutional characteristics (Model C) were included in the analysis.

Facility Educational Conditions. The unexpected results associated with facilities can be better understood by looking at changes in the coefficients, as displayed in Table 4.10. Two findings are noteworthy to discuss. First, all coefficients were significant predictors of test scores in the initial model (Model A); however, the coefficient sizes were smaller than was found in faculty in all fields but Civil Engineering, and the coefficient sign was the opposite of the expectations in "insufficient" condition in Law and in "regular" and "good" conditions in Dentistry.

Second, despite the initial advantage associated with programs with "very good" conditions over most of the other conditions, this advantage was reduced and changed the sign in half of the cases after other variables were added to the regression. Special attention should be given to examining the changes in the coefficients in Model B and Model C. The inclusion of the faculty and curriculum evaluation ratings (Model B) and the institutional characteristics (Model C) in the regression affected not only the size but also the sign of some coefficients. This finding is very important since it shows that facilities had a weak effect on student performance and these results fluctuated when faculty evaluation ratings, curriculum evaluation ratings, and institutional characteristics were added to the regression equation.

Table 4.10 - Regression Coefficients for Facility Educational Conditions

Facility Educational Conditions	Model A Facility Evaluation Ratings	Model B Adding Faculty and Curriculum Evaluation Ratings	Model C Adding Inst. Type and Adm. Control	Model D Adding Student Characte-ristics	Model E Adding Student Involve-ment	Model F Adding Program Selectivity
Administration						
Insufficient	-24.95 ***	-0.25	-14.40 ***	-7.90 *	-8.54 **	8.98 **
Regular	-8.70 ***	13.23 ***	-1.09	1.19	1.59	-1.77
Good	-5.68 ***	2.47	-3.63 **	-0.57	-0.62	2.19
(Reference Group=Very Good Evaluation Outcome)						
Law						
Insufficient	7.11 ***	10.41 ***	14.46 ***	14.37 ***	14.46 ***	5.16 **
Regular	-19.44 ***	-14.54 ***	-8.76 ***	-9.54 ***	-9.85 ***	-11.40 ***
Good	-14.74 ***	-11.14 ***	6.43 ***	5.11 ***	5.68 ***	3.31 *
(Reference Group=Very Good Evaluation Outcome)						
Civil Engineering						
Regular	-64.74 ***	18.03 ***	-3.95	0.65	0.54	2.79
Good	-50.19 ***	8.48 *	-3.58	-2.41	-2.63	-5.25
(Reference Group=Very Good Evaluation Outcome)						
Dentistry						
Insufficient	-31.60 ***	-5.50	-26.21 ***	-20.64 ***	-18.69 ***	-0.30
Regular	12.59 ***	30.21 ***	16.42 ***	18.18 ***	18.85 ***	24.38 ***
Good	21.31 ***	23.86 ***	12.33 **	12.98 **	14.33 ***	16.51 ***
(Reference Group=Very Good Evaluation Outcome)						

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Model A: Only facility evaluation ratings

Model B: Adding faculty and curriculum evaluation ratings

Model C: Adding institutional type (university, university-center, and colleges) and administrative control (federal, state, local, religious/community institutions)

Model D: Adding admission cohort, age, race, income, type of high school, and father's education

Model E: Adding student involvement

Model F: Adding program selectivity

The Effect of Type of High School on Student Performance

According to the results of the regression analysis presented in Table 4.5, students who attended public schools presented higher test scores than students who attended private high school (HS). Since the expectations were to find a negative effect of public high schools on student performance due to the high quality of private secondary education, a block regression was carried out to examine whether the effect associated with type of high schools (e.g., all public HS, both public and private HS, and all private HS) changed after other variables were included in the analyses. In order to achieve that, six blocks of variables were created containing the following information: 1) type of high school, 2) institutional type (university, university-center, and colleges) and administrative control (federal, state, local, religious/community institutions), 3) student's characteristics (admissions cohort, age, race, income, and father's education), 4) evaluation of educational conditions (faculty, curriculum, and facility evaluation ratings), 5) student involvement, and 6) program selectivity. The results of this regression displayed in Table 4.11 revealed two important findings.

First, students who attended public high school (either in full or in part) had an inferior test score compared to their counterparts who attended only private high school in all fields but Dentistry. Interestingly, the big disadvantage was found among students who switched between public and private high schools. In this case, the coefficients associated with attending both public and private high school was reduced by 70% in Civil Engineering and nearly 50% in other fields after institutional characteristics, SES background, and program evaluation ratings were added to the analysis. Regarding

Table 4.11 - Regression Coefficients for Type of High School Attended

Administrative Control	Model A Type of High School Only	Model B Adding Inst. Type and Adm. Control	Model C Adding Student Characte-ristics	Model D Adding Program Evaluation	Model E Adding Student Involve-ment	Model F Adding Program Selectivity
Administration						
All Pub.	-18.67 ***	-15.59 ***	0.35	0.39	0.73	1.37
Public/Private	-30.55 ***	-27.29 ***	-17.40 ***	-16.60 ***	-16.93 ***	-15.89 ***
(Reference Group=Attending all Private High School)						
Law						
All Pub.	-15.16 ***	-11.15 ***	0.07	1.40	1.56	3.68 **
Public/Private	-29.12 ***	-25.18 ***	-15.64 ***	-14.50 ***	-15.21 ***	-13.32 ***
(Reference Group=Attending all Private High School)						
Civil Engineering						
All Pub.	-13.75 ***	-3.51	11.89 ***	12.06 ***	12.05 ***	11.26 ***
Public/Private	-33.64 ***	-24.66 ***	-13.68 ***	-10.10 **	-10.76 **	-10.65 **
(Reference Group=Attending all Private High School)						
Dentistry						
All Pub.	3.47	2.26	7.96 *	8.70 *	8.96 **	8.76 *
Public/Private	-13.35 ***	-9.95 **	-6.14	-6.33	-6.13	-6.04
(Reference Group=Attending all Private High School)						

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Model A: Only type of high school attended (all public, all private, both public and private)

Model B: Adding institutional type (university, university-center, and colleges) and administrative control (federal, state, local, religious/community institutions)

Model C: Adding admission cohort, age, race, income, and father's education

Model D: Adding program evaluation (faculty, curriculum, and facility evaluation ratings)

Model E: Adding student involvement

Model F: Adding program selectivity

students attending only public high school, the results show that the coefficients were drastically reduced after other SES variables were added to the analysis, becoming nearly zero in Administration and Law and positive in Civil Engineering and Dentistry. This result might be related to the high concentration of students from low SES background in public high schools; therefore, the initial disadvantage turned into an advantage when the SES background was controlled.

Second, the expected higher test score of students who attended private high school was not met in Dentistry. Considering that this field presented the highest selectivity level among all fields included in this study, it seems that students admitted to Dentistry had such strong academic preparation that they perform well regardless of the type of high school attended. Besides, the percentage of students who attended public high schools was considerably lower than other fields (as can be seen in Table 3.10), which reinforces the influence of the selectivity of this field in admitting students with stronger academic preparation.

Summary of the Findings

In this chapter, the results of the statistical analyses were presented. The selection of four academic fields enriched the analyses of the relationship between the outcomes of the ENC and ACO, as they offered the opportunity to explore how these relationships differ between academic fields with different characteristics. To conclude this chapter, the most significant findings reported in this chapter are summarized.

One major finding drawn from the descriptive and correlational analyses relates to the supremacy of public institutions in all fields and in almost all aspects analyzed.

Specifically, public institutions were more selective, had higher mean test scores, and had more highly rated faculty and curriculum evaluation ratings in all fields. Separating the aggregation of public and private institutions by the type of administrative control, federally-controlled institutions had higher test scores in all fields, the highest selectivity in Administration and Law, and more highly rated faculty and curriculum evaluation ratings in all fields except Dentistry. State-controlled institutions had the second highest mean test scores in all fields and superior selectivity in Administration, Civil Engineering, and Dentistry. Interestingly, non-profit institutions, followed by for-profit institutions, had the highest facility evaluation ratings in all fields.

Regarding the effect associated with institutional type, the predominance of programs from universities was considerable; such institutions had the highest test score and selectivity in all fields as well as the best faculty and curriculum evaluation ratings in all fields except Dentistry. Another important finding was associated with programs from university-centers which revealed superior facility evaluation ratings in all fields and a higher faculty evaluation rating in Dentistry.

Regression analysis also yielded very important results. First, administrative control (federal, state, local, non-profit, and for-profit) seemed to matter more than institutional type (university vs. non-university) in all fields analyzed. Second, the association between faculty evaluation ratings and student test scores was strong and always significant in all fields examined. Regarding the other program dimensions evaluated by ACO, an initial positive association was found between curriculum and facility educational conditions in some fields; however, the initial association was

drastically reduced, sometimes becoming ninonsignificant, and/or changed the direction of effect after other variables were included in the regression equations. Third, student's SES background was a significant predictor of student test score. The results also showed that income seemed to be a stronger predictor of student performance in the ENC than type of high school attended and father's education. In fact, the results indicated that SES played a different role in the academic fields examined through the strong association with selectivity and/or with test score. Fourth, program selectivity affected test scores directly and indirectly through interacting with institutional type and educational conditions in some fields. Finally, examination of four academic fields with specific particularities offered the opportunity to explore the degree to which some of these findings changed in response to structural differences between fields. In the last chapter, relationships that were not significant will be addressed.

CHAPTER V

ANALYSIS, IMPLICATIONS, AND RECOMMENDATIONS

Introduction

This research project examined the relationship between the outcomes of the evaluation system for undergraduate programs implemented in the Brazilian higher education system. Specifically, this investigation explored the relationship between the educational conditions assessed by a program evaluation (ACO) and student test scores from a national exit exam (ENC). The importance of such an investigation relates to the fact that this evaluation system has embedded in it some assumptions that provide guidance for institutional actions. In order to make appropriate meaning of the findings presented in Chapter 4, this last chapter first summarizes the main findings of this investigation. Next, the two research questions are re-examined, based on the findings of this investigation. Several theoretical lenses are brought into the analysis of the findings in order to help making meaning of the findings. The chapter concludes with implications and recommendations for the Brazilian federal government and for the Brazilian system of higher education, which can potentially be amended and extended to educational domains in a variety of settings around the world.

Revisiting the Main Findings of this Study

This section briefly revisits the main findings with the intention of helping with the re-examination of the research questions. In an attempt to address the main purpose of

this investigation, the relationship between the outcomes of the program evaluation (ACO) and the outcomes of the national exit exam (ENC) was examined by running several statistical analyses that included student characteristics (age, race, admission cohort, variables, socioeconomic status—SES—background, admissions cohort, and student involvement), quality of programs (evaluation of the educational conditions and program selectivity), and institutional characteristics (institutional type and type of administrative control). Five main findings emerged from the results of the analyses.

First, administrative control matters more than institutional type. It is evident that the effects associated with administrative control are much stronger and more meaningful than the effects associated with institutional type. The difference between programs of different types of administrative control was wider than between programs from different types of institutions (as can be seen in Model A in Table 4.6 and Table 4.7). In addition, the coefficients associated with institutional type changed substantially when types of administrative control were included in the regression. The initial positive effect associated with universities turned into a negative effect after the results for administrative control were controlled. On the contrary, the inclusion of institutional type in the analysis did not cause substantial changes in the coefficients associated with administrative control. This means that the coefficients associated with administrative control were much more stable when other variables were included in the regression than the coefficients associated with institutional type. Furthermore, the regression coefficients displayed in Model E in these tables (before controlling for selectivity) show that the advantages of federal and state institutions remained expressive (over two-fifths

of the standard deviation), even after controlling for student's characteristics, quality of programs, and student involvement. The changes caused by the inclusion of selectivity are addressed later.

Drawing from this finding, public institutions demonstrated superior quality (as measured by ACO and ENC) in almost all aspects analyzed. In particular, mean test scores, educational conditions, and selectivity are higher in federal and state institutions than in private institutions. Even when other factors are controlled for, the advantages of these institutions remain impressive. The exception is in Administration, where the effect associated with non-profit institutions on test scores is higher than for all other institutions (including federal and state institutions) after selectivity was controlled. With regard to the institutional type, the results show that although universities had an initial advantage over other types of institutions, this advantage is considerably reduced in Civil Engineering and Dentistry and becomes negative in Administration and Law, after other aspects are included in the regression. This finding by itself is very intriguing, considering the predominance of the university model in Brazilian higher education system.

Second, program selectivity plays a significant role in all fields analyzed. The results show that program selectivity is moderately correlated with test scores and faculty evaluation ratings in all fields analyzed. In addition, selectivity of public institutions is considerably higher than in private institutions in all fields. As a result, it is not surprising that program selectivity is a significant predictor of student test scores in all fields examined. However, the question of whether public colleges and universities have greater

selectivity due to superior "quality" or due to the free tuition policy requires further investigation.

Third, the results associated with the educational conditions assessed by ACO have to be analyzed cautiously and in a specific context. Initial effects associated with higher evaluation ratings in faculty, curriculum, and facilities are consistent with my predictions in all fields. However, this prediction is not met in Law (curriculum) and in Dentistry (curriculum and facilities). Undoubtedly, the results provide evidence for concluding that faculty educational conditions matter more than curriculum and facilities. It does not follow, however, that the other dimensions are not important in the educational process. Indeed, the results suggest that curriculum and facilities are important instruments for promoting higher learning, but that they cannot stand alone and produce better results in the presence of qualified faculty members.

Fourth, the influence of SES background on test scores is among the most significant findings of this investigation. The SES background plays a different role in Administration and Law from that in Civil Engineering and Dentistry. In the biggest fields, in terms of the number of programs and number of students in Brazilian Higher Education (Administration and Law), SES is a significant predictor of student test scores. In these fields, a higher family income corresponded directly a higher test score, even when all other variables are incorporated into the regression equations. Father's education level is a significant predictor of test scores only when it is included in the regression alone; however, its effect is considerably reduced after controlling for other variables, becoming nearly zero in Law and minimal in Administration. Additionally, when

analyzing the effect of the type of high school attended, the initial positive effect on test scores associated with having attended a private high school drops dramatically and becomes negative (but nearly zero) once family income is added to the regression. In Civil Engineering and Dentistry, two measures related to the SES background—income and father's education level—are insignificant predictors of test scores, after all variables are included in the regression.

Lastly, the inclusion of institutional characteristics, educational conditions, and program selectivity affect the size and in some cases the direction of the effect of other variables, making the interpretation of the findings challenging. In order to develop a meaningful understanding of the findings, it is imperative that one have a clear understanding of the context in which the program assessment (ACO) and the national exit exam (ENC) are situated and of how different institutional settings affect the evaluation outcomes. This analysis is presented later in this chapter, when the findings are analyzed.

Revisiting the Purpose and Research Questions

The purpose of this investigation was to examine the degree to which student performance at the ENC is related to program characteristics assessed by ACO. In order to achieve this aim, two research questions were posited.

To restate the first research question: *What are the relationships, if any, among the mean test score (ENC), faculty evaluation rating, curriculum evaluation rating, facility evaluation rating, program selectivity, and institutional characteristics (institutional type and administrative control)?* The results of the correlation analysis

indicate that there are important relationships among these variables. Mean test scores are positively associated with program selectivity and faculty evaluation ratings in all fields examined. However, the association between test scores and curriculum and facilities (the other educational conditions) does not meet the prediction---they are not correlated in Law (since the coefficients are nonsignificant and practically zero) and are positively correlated but not significant in Dentistry. Regarding the correlation among the educational conditions evaluated by ACO, the results show that all three dimensions are positively correlated, showing the strongest correlation in Civil Engineering and the weakest in Administration. Lastly, the results of the descriptive analysis show that federal and state institutions, followed by non-profit institutions, have higher test scores, program selectivity, and faculty evaluation ratings than other types of institutions, those being the locally controlled and for-profit institutions. This finding suggests the existence of an association among the institutional type (specifically, the administrative control), quality or programs (as assessed by the ACO), and student performance factors at the ENC.

To restate the second research question: *To what extent is the student test score on the ENC associated with faculty evaluation rating, curriculum evaluation rating, and facility evaluation rating, while holding constant SES background, individual characteristics (race and age), program selectivity, and institutional characteristics (institutional type and administrative control)?* The findings demonstrate, in general, that educational conditions as evaluated by ACO have an initial positive and significant effect on test scores in almost all fields examined. As expected, a higher faculty evaluation rating corresponds to a higher test score in all fields, even after all variables are added to

the model. However, the results associated with curriculum and facilities are mixed and in some cases contradictory. Curriculum evaluation rating proves to be a significant predictor of test score only in Administration and Civil Engineering, failing to present any effect on test scores in Law, and having an unexpected result in Dentistry. The effect associated with programs with "regular" and "good" curriculum evaluation ratings in Dentistry is greater than the effect associated with programs with "very good" evaluation ratings. With regard to facilities, the results indicate that better facilities do not lead to higher test scores in all fields, after controlling for other variables. The results of several regression analyses indicate the existence of important differences between the fields examined in terms of the association between the educational conditions and test scores, as well as the influence of program selectivity, type of administrative control, and student SES background on test scores.

Making Meaning of the Findings

Having summarized the main findings and having related them to my two research questions, this section presents an exploration of and an explanation of the findings using several theoretical lenses. As the findings are discussed, one or more theoretical frameworks related to the topic in question are brought into the analysis as a means for reaching an understanding of the findings. The results associated with institutional characteristics (administrative control and institutional type) are discussed first, due to their influential role throughout the remaining findings. Next, the results associated with educational conditions and SES background are presented.

The Influence of Institutional Characteristics on the Outcomes of the Evaluation System

Overall, the quality of programs from federal and state universities is superior to the quality of programs from local and private institutions. These institutions had the highest mean test scores, the highest evaluation ratings with respect to faculty and curriculum, and superior selectivity in three of the fields examined in this investigation—Law, Civil Engineering, and Dentistry. This superiority makes it relevant to inquire as to why one particular sector of the higher education system has better educational inputs and educational conditions than the other sectors.

The literature regarding institutional stratification provides the basis with which to examine the advantages that some institutions had in almost all of the dimensions examined in this investigation. Institutions are stratified based on the prestige, wealth, and power they are able to acquire, which might be a result of the influence of the market or the actions promoted by governments (Trow, 1984). In the Brazilian higher education system, instead of being influenced by market principles, the institutional hierarchy of the higher education system seems to be the most strongly influenced by the actions promoted by federal and some state governments.

A number of aspects might be associated with the advantage these institutions have with regard to the outcomes of the evaluation system for undergraduate programs. First, the Brazilian higher education system is strongly regulated by federal legislation—which provides the legal framework for federal and private institutions—and by state legislation—which provides the legal framework for state and local institutions. An example of this regulation is the legal framework that created this evaluation system. The

evaluation instruments (ACO and ENC) were created by federal legislation, and all institutions of higher education have to participate in this evaluation process.

Second, federal universities and some state universities were idealized based on a model of research universities that guided the expansion of the Brazilian higher education system in previous decades. Universities, under federal legislation, are institutions of higher education which have to integrate teaching, research, and community services. Since the money available for financing the development of research comes primarily from the federal government and there are strong restrictions on allocating public resources to private institutions, federal and some state universities maintain a quasi-monopoly of graduate education and research.

Third, governments are responsible for nearly all of the expenses of public colleges and universities. Instead of having to generate resources from tuition, which could create a problem for raising more resources, federal institutions were able to increase the budget allocated to higher education in previous decades by negotiating directly with the federal government. Up to the mid-1990s, public governments (especially the federal and some state governments) invested heavily in the infrastructure necessary to develop research and to provide graduate education. As a result, public institutions were able to pay higher faculty salaries and built the infrastructure required to develop research; this in turn attracted faculty members with high academic credentials who were willing to develop research and participate in graduate programs.

However, public appropriations for higher education have not increased in the same proportion as enrollment growth during the last decade, which has created a marked

financial crisis for public universities. One factor that has influenced this policy change relates to the financial constraints faced by governments in the last decade (especially the federal government). In fact, reduction of public appropriations for higher education has also been observed in U.S. higher education as well as in many other countries. It is important, however, to point out that neo-liberal policies defended by international organizations, such as World Bank and Inter-American Development Bank (IDB), call for reducing the presence of governments in higher education in Latin America. Despite the financial crisis faced by public institutions, public investment in higher education is clearly working in Brazil as public institutions are still able to provide education at a higher quality than private institutions.

Fourth, the free tuition policy, along with other advantages some elite public institutions have, has created more leverage for these institutions. The influence of the free tuition policy on the quality of Brazilian colleges and university opposes the conventional wisdom present in the U.S. system. In the Brazilian case, selectivity along with the high quality of public universities becomes a key aspect of the stratification of this higher education system. The extreme difference in selectivity between public and private institutions raises concern about the influence on the results of the test scores in the exit exam, a concern expressed by Astin (1985) and Ewell (2001). Although the data do not allow me to investigate the effect of student selection biases, it is very likely that the advantages offered by federal and state universities in terms of the free tuition along with the superior educational quality (demonstrated by the ACO outcomes) create initial differences between students admitted to elite public institutions and students admitted to

non-elite institutions (usually called selection bias). Therefore, part of the advantages associated with being a federal or state institution is likely due to initial differences in the selection process, as the regression results suggest. In fact, Astin and Lee (2003) address this problem in what they called one-shot assessment that lacks a pretest, arguing that the assessment outcomes do not reflect the actual influence of colleges on students. They state that institutions with high scores should not necessarily take credit for over-performing and that institutions with low scores should not necessarily be blamed for under-performing because the assessment results are likely to reflect initial differences between students.

The value-added approach defended by Astin (1991) and Soares, Ribeiro, and Castro (2001) could overcome this limitation by focusing on how much a certain institution was able to help students learn. However, the lack of a measure of pre-college academic preparation available at the national level presents the most significant limitation of this evaluation system, since it is not possible to examine the value added by different types of colleges. Although nearly all Brazilian colleges and universities have their own entrance exams used to select freshmen, these exams could not be used in a study at the national level for two reasons. First, the results of this exam are not available through any national databases. Second, there is no mechanism that could link the test score from the *vestibular* and the test score from the ENC, due to the confidentiality of the ENC test scores.

In sum, the mechanisms under which institutional stratification is promoted in the Brazilian higher education system seem to strongly affect the evaluation outcomes of the

ENC and ACO. It is important to recall that the way in which Brazilian colleges and universities are stratified is very different from the U.S. system. The allocation of resources, the quasi-monopoly of research and graduate education, as well as the free tuition policy, have created an overwhelming advantage for public institutions over private institutions, especially small colleges. This advantage seemed to be translated into higher faculty and curriculum evaluation ratings, higher selectivity, and higher mean test scores, creating a virtuous cycle for these institutions as defined by Merton (1973). Private institutions, which charge tuition, have to compete in this adverse environment. The strong reliance on student tuition, the low contribution level of private corporations to higher education, the modest federal loan program, and the lack of funds such as endowments make it difficult for private institutions to increase expenses in order to compete with public institutions. Elite private institutions do exist in Brazil and they have been able to compete with public institutions in terms of the quality of education. However, they have to be seen as elite institutions that also benefit from the prestige and wealth they were able to accumulate over the years.

The Influence of Educational Conditions on Student Test Scores

The basic assumption embedded in this evaluation system is that better educational conditions (ACO) are assumed to be associated with higher test scores (ENC). The lack of evidence concerning this association provided the motivation for conducting this investigation. The need to explore such an important relationship is explained by the fact that if scientific evidence is found in favor of this assumption,

educational conditions (e.g., faculty, curriculum, and facilities) should be seen by educators and policymakers as factors that could be manipulated by colleges and universities to promote higher student learning. However, this association was not found in all educational conditions evaluated, requiring a careful examination of the context of this finding. In order to consider the context in which this result was found, the mean test scores, the initial coefficient, and the final coefficient for each outcome of the educational dimensions assessed by ACO are displayed in Table 5.1. This table was included here to show that the results of the final model associated with the educational conditions can be explained by one of the following situations: 1) the results met the prediction; 2) the results can be explained by the differences in mean test scores between groups; and 3) the results indicate the existence of an interaction between educational conditions, institutional type, and selectivity. The examination of the influence of the faculty evaluation rating on student performance is presented first, followed by the influence of curriculum and facilities.

Faculty Evaluation Ratings. Faculty is the most critical resource for success in college. In general, a higher faculty evaluation rating corresponds directly to a higher test score, even when all other variables are incorporated in the regression equations. In all fields examined, when the faculty received "very good" ratings, the test scores correlated significantly with this rating. This positive association supports the important role played by faculty in the educational process.

The evidence strongly indicates that the faculty should be seen as a valuable resource that certainly influences student learning inside and outside the classroom. In

this context, faculty must be seen as the most important resource for promoting success. In fact, a great deal of research in higher education has addressed the role of faculty in fostering student learning, mostly inside the classroom. For instance, student learning can be enhanced, according to Chickering and Gamson (1991), through the following academic good practices: student-faculty integration, cooperation among students, active learning, prompt feedback, emphasis time on task, communicating high expectations, and respect to diversity of talents and different ways of learning. In addition, Cross (1991) suggests that good teaching is associated with the importance of knowledge of the subject, the ability to provide clear explanations, being accessible to students, and providing an organizational plan, among others.

It is also interesting to analyze faculty would serve as a resource to students. The criteria used to assess faculty in the ACO evaluation include, among others, faculty credentials (e.g., B.A., M.A., and Ph.D.), publications, and participation in research and scientific events as criteria for evaluating faculty (some of these criteria as also utilized in the U.S. system with the U.S. News & World Report Ranking). In this context, faculty shouldbe seen as mentors to students as well as network providers under which other resources are properly utilized in the educational process to facilitate higher leaning. Faculty who earn an advanced degree abroad, who participate in scientific events, and who conduct research might be much more prepared to mentor students who want to entering graduate education. Although this hypothesis seems plausible from the findings of this investigation, the data do not allow us to understand exactly which of these

characteristics are more related to student learning and how they foster student learning. Further research must be done in order to explore this relationship.

However, the association between faculty evaluation ratings (ACO) and test scores (ENC) might seem to reveal that both assessment instruments were designed based on some aspects that do not influence student learning. For instance, it might be the case that both evaluation instruments (ACO and ENC) have biases against less prestigious institutions and/or students from low socioeconomic (SES) families. Since these characteristics are mostly present in research universities, the differentiation of programs with "good" educational conditions from programs with "very good" conditions might be inappropriate for investigating student performance in all types of institutions. A similar concern is present in the literature with regard to the utilization of certain abilities and skills more commonly found among students from high socioeconomic (SES) backgrounds to discriminate student performance in such tests as the ENC (Bogue & Sounders, 1992; Flaugher, 1974; Kohn, 2000; Shepard, 1981).

Curriculum and Facility Evaluation Ratings. Unlike the findings associated with faculty assessment ratings, the effect of curriculum and facility on student test scores is mixed and confusing. The prediction that a higher curriculum evaluation rating (as well as a higher facility evaluation rating) would correspond directly to a higher test score in all fields is not met, even when each of these dimensions is included in the regression equations alone. This finding might lead to incorrect inferences about the influence of curriculum and facilities on test scores if further analysis is not carried out to explain the reasons for such unexpected results.

The effect associated with faculty in all fields examined and the effect associated with curriculum in Administration and Civil Engineering meet the predictions, as displayed in Appendix B.3. In these cases, a higher evaluation rating corresponds directly to a higher test score, even when all variables are included in the regression equations. However, the findings associated with curriculum in Law and Dentistry as well as facilities in all fields do not support the main assumption of this evaluation, which is that superior educational conditions will be associated with higher test scores.

The reason for the "unexpected" finding regarding curriculum and facilities in Law and Dentistry is revealed when the mean test scores in each group are examined. Programs with "regular" and "good" evaluation ratings in curriculum and facility in Dentistry have higher mean test scores than programs with "very good" conditions, which is consistent with the final coefficients. With respect to the effect associated with curriculum and facilities in Law, the mean test scores shows that programs with "regular" curriculum evaluation rating have the highest mean test scores, followed by programs with "very good", "good", and "insufficient" evaluation ratings (coefficients showed to follow the same order).

Lastly, the results associated with facility evaluation ratings in Administration and Civil Engineering should be viewed carefully. Despite the fact that mean test scores show the predicted relationship with the facility evaluation ratings (as can be seen in the "initial" coefficient), the results associated with the coefficients from the final model raise a concern about the existence of an interaction between facility and other variables included in the regression analysis. In fact, facility evaluation rating is moderately

correlated with curriculum and faculty and all evaluation ratings also correlate with student test scores. This suggests that faculty, curriculum, and facilities share a significant amount of predictability with respect to the test score. Therefore, the coefficients associated with facility evaluation ratings drop considerably and some coefficients change the direction of effect when all variables are added to the regression equation, especially institutional type and program selectivity. This result is not surprising since faculty evaluation ratings, program selectivity, and institutional characteristics have a much stronger association with test scores than do faculty evaluation ratings.

These unexpected results associated with curriculum and facility can be better understood considering that institutional compliance in one of the incentives embedded in this evaluation policy. As discussed in the first chapter, the outcome of the evaluation of educational conditions depends on the level of compliance with standards set for each academic field, representing the extent to which undergraduate programs have achieved or exceeded the specified "quality" standards. The concern raised by Burke and Minassians (2002) and Richard (1988) with regard to the development of institutional efforts that focus more on compliance rather than on improvement seems to be present in the results of the program evaluation (ACO). Although this was not explored during this investigation, it seems very plausible that institutions of higher education, especially non-elite institutions, have developed strategic actions based on the incentives embedded in this evaluation policy. Because it seems to be much more expensive to invest in faculty to raise the evaluation ratings, institutions may have focused their efforts on increasing the

level of compliance in the dimensions of curriculum and facilities, which requires less investment in the long run. As a result, institutions whose faculty evaluation ratings and mean test scores were among the lowest presented the highest facility evaluation ratings. This institutional effort to raise the facility evaluation ratings might be also seen as a strategic action taken to increase only one educational condition, thereby reducing the burn associated with having "poor" evaluation ratings in all dimensions assessed. At the same time it is important to offer students adequate educational facilities, this strategy by itself has shown to be inadequate to enhance student leaning, for the reasons summarized herein.

In sum, the analysis of the effect associated with all three educational conditions assessed by ACO seems to suggest that faculty matter most in the teaching process; however, curriculum and facilities can also make a difference in this process. Undoubtedly, the analysis of the results associated with faculty evaluation ratings presents the most consistent results among all educational dimensions assessed. However, the effect of curriculum and facilities on test scores seems to depend on a much broader perspective that involves considering the influence of faculty evaluation ratings, institutional settings, and program selectivity. Returning to the discussion about the "mixed" and "confusing" effect associated with curriculum and facilities, it seems possible to conclude that "good" and "very good" evaluations of these educational conditions do not necessarily lead to a propitious environment for promoting student learning. In other words, policymakers and educators should see curriculum and facilities as an indispensable means to promote education at a high "quality" standard; however,

the success of the teaching process requires faculty members who are well prepared to foster student learning, which requires the existence of a good curricular structure and well-equipped educational facilities.

The Influence of Socioeconomic (SES) Background on Student Test Scores

The effect of socioeconomic (SES) background on test score in the fields examined is mixed. The theory of social reproduction (Bourdieu, 1973) provides the basis for examining the effect associated with socioeconomic (SES) background. In particular, the notion of cultural capital offers a valuable lens to understand why test scores seem to be related to SES.

Considering only the effect associated with SES in the full regression model (which included all variables), socioeconomic (SES) background is a significant predictor of student performance in Administration and Law, but an insignificant predictor in Civil Engineering and Dentistry. Despite this "unexpected" finding, it seems that the mechanisms under which the SES influenced test scores in these academic fields differed. While the influence of SES directly translates into the final coefficients in Administration and Law, its influence seems to be indirect and related to the high program selectivity in Dentistry.

In Administration and Law, socioeconomic (SES) background is a significant predictor of test scores. Programs from these fields enroll a large number of students from different socioeconomic (SES) backgrounds. Even in Law, where the average admission ratio is close to ten percent, a large number of students from low SES

backgrounds are admitted to college. The analysis of the results is straightforward, indicating that strong evidence exists to conclude that students from high SES backgrounds perform better in the national exit exam (ENC) than their counterparts from low SES backgrounds, even after controlling for other variables.

This positive association might be due to two aspects. First, students from high SES backgrounds generally come to college with a much stronger academic preparation than students from low SES background. Furthermore, these students, according to the theory of social reproduction, have acquired the cultural capital necessary to succeed in the educational system. As a result, they have not only a better academic preparation but also a high congruence between their own values and the values promoted by the educational system (Dumais, 2002). These aspects might create an advantage for high SES students and may be responsible for the effect associated with socioeconomic (SES) background. This process, therefore, enforces the perpetuation of social inequalities (McDonough, 1994; Nogueira & Nogueira, 2002).

An alternative explanation is related to the existence of biases against students from low SES background in the formulation of the national exit exam (ENC). In this case, designers of norm-referenced tests (NRTs) make use of certain characteristics mostly present among students from high SES backgrounds to discriminate student performance. As a result, the differences in test scores might be associated with socioeconomic (SES) background rather than with what is actually learned in college (Kohn, 2000; Shepard, 1981).

Another important finding revealed from the analysis of programs in Administration and Law relates to the influence of family income on test scores. In these fields, income is shown to have stronger effect on test scores than other variables associated with SES. It might be the case that attending a private high school provides better preparation for college than attending a public high school. Having parents with a college degree might represent an advantage for high SES students, since parents not only have more preparation to help their kids throughout their elementary and secondary education, but also have higher expectations regarding attaining a college degree. Significantly, having a certain income level enables families not only to enroll their children in private schools but also to attend cultural activities that help children acquire the cultural capital necessary to succeed in the educational system. It seems to be the case that once the family income and father education level are considered, the difference in having attended a public or a public high school is considerably reduced. Finally, the Brazilian educational system has undergone significant transformations in recent decades. Access to higher education used to be strongly limited in the past so that only a small proportion of families had parents with college degree. However, the trend in economic development has promoted an upward social mobility, which has resulted in more parents with a higher income level than with a college degree. For example, data from this investigation show that sixty percent of parents with incomes in the highest level also have a college degree.

The analysis of the results associated with Civil Engineering is more complex. Income is an insignificant predictor of test score in the full regression model. Even when

the effects of SES variables on test scores are individually tested, only the effect associated with the top two income levels is significant. This finding reveals that students from middle and low SES background do not have significantly different test scores. One possible explanation might be related to some differences between students who enroll in Civil Engineering, compared with other fields. Despite the fact that the average selectivity is similar to the selectivity in Administration, students who apply for Civil Engineering might have a better academic preparation than other students. However, this hypothesis could not be explored because of the lack of a measure of pre-college preparation.

With regard to Dentistry, the nonsignificant effect associated with SES background might be due to the selectivity of the field. With an admission rate of only 5.7 percent, students who were admitted to a program in this field should theoretically have a stronger academic preparation than students admitted to other fields. In addition, given the small number of spaces offered in this field by all institutions, compared with Administration and Law, low SES students enrolled in a smaller percentage than in other fields. As a result, the SES profile of students enrolled in this field is considerably higher than that of other fields: they are younger, more likely to have attended a private high school, and had families with high SES background. Because of the selectivity and the small number of spaces for freshmen, students from low SES background did not receive lower test scores than students from other SES backgrounds.

Implications and Recommendations

The findings of this investigation lead to several implications, especially for the federal government. Even though this research did not seek to examine actions taken by colleges and universities in response to the implementation of this evaluation system, the main interest of colleges and universities should be whether the evidence derived from this investigation could be used to promote actions to enhance student learning. However, it is important to note that this evaluation system presents other incentives for institutions, such as complying with educational standards set by national committees. For that reason, it is extremely important to investigate how institutions have acted under the framework imposed by this evaluation because it is possible to raise certain educational standards without necessarily enhancing student learning.

Most of the implications refer directly to the federal government because the main purpose of this investigation was to explore the underlying assumptions of this educational policy, which was designed to address quality in higher education. Therefore, the implications have to be necessarily related to the evidence produced to either support or reject the underlying assumptions of the current evaluation system for undergraduate programs. From the policymaking standpoint, the findings are important because they provide policymakers and educators with the guidance required to develop institutional actions aiming to deal with quality of education. Furthermore, awareness about the possible biases incorporated in the evaluation instruments, as well as the lack of evidence for some of the assumptions that underlie this evaluation policy, should be as important as the evidence concerning the effectiveness of this evaluation system. This offers the

opportunity to promote changes in the evaluation system, as well as to identify the circumstances in which the evaluation outcomes should be used in order to make appropriate use of them.

Four specific implications for the federal government as well as for institutions of higher education are presented herein. First, investment in faculty pays. Enough evidence from this investigation exists to conclude that faculty is the most critical resource necessary to promote student learning. Undoubtedly, highly rated faculty yields better results than highly rated curriculum and facilities. This point of view is clearly supported by the fact that private institutions have invested a lot in facilities; however, this investment seems not to directly translate into higher test scores. Additional support can be found from the fact that lack of investment in facilities in public institutions indicates that high quality education does require qualified faculty but not necessarily high quality facilities. Therefore, the development of institutional actions designed to enhance the quality of faculty should be the main focus of the public policy but policymakers should be aware that curriculum and facilities can make a real difference with qualified faculty members.

Second, the evaluation outcomes (ENC and ACO) must be used cautiously, with the understanding that elite institutions are likely to have better educational outcomes than non-elite institutions. In fact, enough evidence exists to conclude that elite institutions (especially federal and state research institutions) provide education at a higher quality, which might be influenced by the public investment in education, the quality of their students, the quality of their programs (more specifically faculty), or any

combination of these factors. For these reasons, the results of this evaluation system should be used, especially considering that non-elite institutions are unlikely to achieve higher evaluation outcomes due to some of their institutional characteristics and the type of students they enroll. One example of the influence of institutional characteristics on the evaluation outcomes can be found in the evaluation of faculty by ACO. The findings showed that federal and state institutions have higher faculty evaluation ratings than other kinds of institutions. The criteria utilized to determine this rating include faculty credentials (M.A. and Ph.D.), dedication to the program (full-time versus part-time), faculty publications, and participation in scientific events, which are mostly present in programs from research institutions (especially federal and some state institutions). However, mixing in single measure (faculty educational conditions) characteristics directly associated with promoting teaching at a high quality level with other characteristics required for promoting high quality research makes the analysis of the relationship between faculty educational conditions and student performance complex.

Therefore, further research should break down the composite ratings produced by the ACO as the final evaluation ratings in order to examine which program characteristics are required to *enhance* student learning or whether they are only *associated with* student learning. By doing so, it would be possible to provide much stronger evidence about, for instance, specific characteristics of faculty that seem to be responsible for enhancing student success in college. In addition, the differences between institutions and how student success should be fostered would be very clear for policymakers. Even if no evidence can be found to support the positive effect of faculty on enhancing student

learning, a different measure associated with development of research should be created since quality in higher education cannot be seen only as a function of student learning.

Third, policymakers must be educated to interpret the results of the exit exam (ENC) within a certain context because they are extremely sensitive to student socioeconomic (SES) background and program selectivity. For instance, the influence of SES and program selectivity on student performance on the ENC presents a real challenge for the use of the results of this test as a measure of program effectiveness. Students from high SES background and students who enrolled in selective programs also had higher test scores than students from low SES background and students enrolled in less selective programs. Because of the lack of a measure of pre-college preparation that could be used to control for previous difference in academic preparation, it is not possible to positively conclude that the performance on the ENC exam was entirely due to institutional efforts or to previous individual differences. Therefore, the result of the national exit exam (ENC) should not be used as a proxy for program effectiveness without considering pre-college academic preparation of students and the selectivity of the program they enrolled. While program selectivity can be controlled for, as it was in this research, the lack of a measure for pre-college preparation imposes a real limitation for this evaluation system.

In order to overcome this limitation, the federal government should include a measure related to pre-college academic preparation in order to control for pre-college differences. Such a measure could be the test score from ENEM, a national achievement test taken by high school students. Since this test is not mandatory, some students do not

take the test. One option would be to make this test mandatory for all students applying for college. Another option that has been discussed now in Brazil includes a kind of pre-test applied to a sample of college students in their freshmen year. The availability of a pre-test along with a post-test could make it possible to infer the value added by colleges, reducing considerably the problem associated with selection bias currently present in this evaluation system. Furthermore, the quality or effectiveness of a given program would be assessed by the result of its effort in helping students to learn, regardless of the initial condition in which they were admitted to college. Using this design, institutions of all kinds could benefit from the results of this type of evaluation.

Fourth, the focus of the current evaluation system on the program level instead of on the classroom level has important consequences for the decision-making process. Due to the fact that the evaluation outcomes consolidated at the program level have been used by the federal government in the accreditation process of undergraduate programs (Ferrer, 2001; Gomes, 2002) and by the media to rank undergraduate programs and higher education institutions (Segenreich, 2000), strong pressure is created to promote changes that will result in better evaluation outcomes. This pressure leads institutions to concentrate their efforts on promoting changes in the dimensions assessed by ACO, as some authors have discussed (Caldeira, Kraemer, & Vasconcelos, 2000; Freitas, 2000; Martins, 2000; Silva, 2001). This might have created a compliance culture rather than a culture necessary for improving the teaching-learning process.

In order to create an environment propitious to enhancing the quality of education and educational outcomes, it is necessary to move toward implementing assessments at

the classroom level, as Nordval and Braxton (1996) have argued. The importance of having assessments at that level is justified by the fact that high faculty credentials, a considerable number of publications, and participation in research and scientific events do not guarantee the efficacy of the teaching process, although they certainly represent an important dimension of the educational quality. Therefore, it is equally necessary to assess the teaching-learning process inside the classroom. Work by Chickering and Gamson (1991) entitled "Seven Principles for Good Practice in Undergraduate Education" provides an initial framework for investigating some aspects that might actually enhance student learning rather than simply raising educational standards. Although it is very unlikely that a national evaluation policy will incorporate assessments at the classroom level, policymakers have to be aware of the need to develop such assessments, probably coordinated by internal evaluation committees. However, in order to implement successful evaluations at the classroom level, it is necessary to have the institutional capacity to carry out these assessments, federal incentives to fund these types of initiative, and mechanisms to report successful experiences.

To conclude, three general principles are helpful in informing public policymakers. First, investment in public education pays, as the educational outcomes of public institutions were significantly higher than those from private institutions. Second, faculty are the most critical resource for student success. Therefore, investment in faculty is required to provide high quality education, which in turn does not require high quality facilities. It does not follow, however, that curriculum and facilities do not influence student learning. They do matter but in the presence of qualified faculty. Third,

investment in graduate education is extremely necessary as the means to prepare qualified faculty to teach in all institution types, especially in this period of significant enrollment expansion within Brazilian higher education. The growth of graduate education would provide institutions of higher education with more qualified faculty, thereby reduce the existing gap between elite and non-elite institutions.

Conclusion

The concern about quality has definitively reached the national agenda of the Brazilian higher education system since the implementation of the national evaluation system for undergraduate programs. This evaluation system, as an educational policy, has embedded some assumptions related to the quality of education, as well as to how this quality should be promoted. Through these assumptions, institutions of higher education are likely to guide their actions, aiming to meet the standards established by the federal government. In an attempt to explore the underlying assumptions of this evaluation system, this investigation investigated the relationship between the educational conditions of undergraduate programs assessed by ACO and student test scores from the national exit exam (ENC). It is very clear that elite institutions (especially research universities), as well as high SES students, benefit from this evaluation system through receiving better evaluation outcomes and higher test scores, respectively. The underlying assumption about the direct relationship between educational conditions assessed by ACO and student performance on the ENC is partially supported by the study results. The findings also reveal some concerns about the existence of bias against non-elite institutions and low SES students, as well as against the establishment of an institutional culture that

focuses more on compliance than on improvement. The implications of these findings for institutions of higher education and for the federal government were discussed, along with some recommendations necessary to overcome some problems.

APPENDICES

APPENDIX A – DEFINITION OF TYPES OF INSTITUTIONS AND ADMINISTRATIVE CONTROL

Institutional Types	
Universities	These institutions typically offer a wide range of baccalaureate programs as well as graduate programs, are committed to research and extension services, and are autonomous to open new programs without previous authorization from the federal government (includes also specialized universities that are institutions with major emphasis on one academic field)
University-Centers	These institutions typically offer a wide range of baccalaureate programs in many academic fields and are autonomous to open new programs without previous authorization from the federal government (includes also specialized university-centers that are institutions with major emphasis on one academic field)
Colleges	These institutions typically are small institutions that offer baccalaureate programs in one of few academic fields but do not have autonomy to open new programs without previous authorization from the federal (or state governments)
Administrative Control	
Public Institutions	The institutions are controlled and funded by the federal, a state, or a local government
Federal Institutions	They are public institutions controlled and funded by the federal government
State Institutions	They are public institutions controlled and funded by a state government
Local Institutions	They are public institutions controlled and funded by a local government
Private Institutions	The institutions can be non-profit institutions or for-profit institutions
Non-profit Institutions	They are private institutions that are not allowed to profit from their educational activities and have some benefits such as tax-exempt. They can be community institutions, religious-oriented institutions, or philanthropic institutions.
Community-Oriented Institutions	They are non-profit private institutions controlled by one or more individuals and must have some members from the community in the Board of Regents
Religious-Oriented Institutions	They are non-profit private institutions with religious or ideological orientation
For-profit Institutions	They are private institutions that are allowed to profit from their educational activities; however, they operate under the legislation that controls all other private organizations and have to pay all taxes like other private organizations

Appendix B.1.1 -Mean, Standard Deviation, Minimum, Maximum of Student-Level Variables - Program: Administration

Variable	N	Mean	SD	Min.	Max.
Student-Level Variables (N=30776)					
Student Test Score (ENC)					
Test Score (Multiple-Choice)		37.52	11.68	0.00	85.00
Test Score (Open-Question Format)		26.24	14.91	0.00	90.00
Test Score (All Questions)		31.90	11.14	1.30	76.30
Standardized Test Score		500.00	100.00	225.21	898.62
Admission Cohort					
Admitted in 1995 or Before	3,451	0.11	0.32	0.00	1.00
Admitted in 1996	3,615	0.12	0.32	0.00	1.00
Admitted in 1997	9,430	0.31	0.46	0.00	1.00
Admitted in 1998 or Later	14,280	0.46	0.50	0.00	1.00
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	16,194	0.53	0.50	0.00	1.00
Age: 25-30 Years Old	8,974	0.29	0.45	0.00	1.00
Age: Older than 30 Years Old	5,608	0.18	0.39	0.00	1.00
Race					
White	25,441	0.83	0.38	0.00	1.00
Black/Mulato	4,067	0.13	0.34	0.00	1.00
Asian	1,038	0.03	0.18	0.00	1.00
Native	224	0.01	0.09	0.00	1.00
SES Related Variables					
Family Income					
Income: <= R$ 600	1,228	0.04	0.20	0.00	1.00
Income: R$ 601 to R$ 2.000	9,779	0.32	0.47	0.00	1.00
Income: R$ 2.001 to R$ 4.000	10,414	0.34	0.47	0.00	1.00
Income: R$ 4.001 to R$ 10.000	7,147	0.23	0.42	0.00	1.00
Income: >= R$ 10.001	2,208	0.07	0.26	0.00	1.00
Parental Education					
Father Having HS Degree	7,097	0.23	0.42	0.00	1.00
Father Having College Degree	8,504	0.28	0.45	0.00	1.00
Mother Having HS Degree	8,377	0.27	0.45	0.00	1.00
Mother Having College Degree	6,786	0.22	0.41	0.00	1.00
Type of High Scholl Attended					
All Private HS	13,310	0.43	0.50	0.00	1.00
All Public HS	12,752	0.41	0.49	0.00	1.00
Some Public / Some Private	4,714	0.15	0.36	0.00	1.00
Involvement in College Activities					
Student Involvement		3.69	1.37	0.00	6.00

Appendix B.1.2 -Mean, Standard Deviation, Minimum, Maximum of Student-Level Variables - Program: Law

Variable	N	Mean	SD	Min.	Max.
Student-Level Variables (N=34834)					
Student Test Score (ENC)					
Test Score (Multiple-Choice)		43.22	14.69	0.00	97.50
Test Score (Open-Question Format)		37.99	18.78	0.00	100.00
Test Score (All Questions)		40.61	13.88	1.25	90.00
Standardized Test Score		500.00	100.00	216.35	855.97
Admission Cohort					
Admitted in 1995 or Before	2,940	0.08	0.28	0.00	1.00
Admitted in 1996	7,943	0.23	0.42	0.00	1.00
Admitted in 1997	23,130	0.66	0.47	0.00	1.00
Admitted in 1998 or Later	821	0.02	0.15	0.00	1.00
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	19,411	0.56	0.50	0.00	1.00
Age: 25-30 Years Old	8,318	0.24	0.43	0.00	1.00
Age: Older than 30 Years Old	7,105	0.20	0.40	0.00	1.00
Race					
White	29,272	0.84	0.37	0.00	1.00
Black/Mulato	4,516	0.13	0.34	0.00	1.00
Asian	727	0.02	0.14	0.00	1.00
Native	300	0.01	0.09	0.00	1.00
SES Related Variables					
Family Income					
Income: <= R$ 600	1,530	0.04	0.20	0.00	1.00
Income: R$ 601 to R$ 2.000	9,700	0.28	0.45	0.00	1.00
Income: R$ 2.001 to R$ 4.000	11,239	0.32	0.47	0.00	1.00
Income: R$ 4.001 to R$ 10.000	9,668	0.28	0.45	0.00	1.00
Income: >= R$ 10.001	2,697	0.08	0.27	0.00	1.00
Parental Education					
Father Having HS Degree	7,555	0.22	0.41	0.00	1.00
Father Having College Degree	14,687	0.42	0.49	0.00	1.00
Mother Having HS Degree	9,926	0.28	0.45	0.00	1.00
Mother Having College Degree	11,888	0.34	0.47	0.00	1.00
Type of High Scholl Attended					
All Private HS	18,519	0.53	0.50	0.00	1.00
All Public HS	10,739	0.31	0.46	0.00	1.00
Some Public / Some Private	5,576	0.16	0.37	0.00	1.00
Involvement in College Activities					
Student Involvement		4.07	1.33	0.00	6.00

Appendix B.1.3 -Mean, Standard Deviation, Minimum, Maximum of Student-Level Variables - Program: Civil Engineering

Variable	N	Mean	SD	Min.	Max.
Student-Level Variables (N=4880)					
Student Test Score (ENC)					
Test Score (Multiple-Choice)		-	-	0.00	0.00
Test Score (Open-Question Format)		29.08	15.92	1.00	92.00
Test Score (All Questions)		29.08	15.92	1.00	92.00
Standardized Test Score		500.00	100.00	323.56	895.33
Admission Cohort					
Admitted in 1995 or Before	1,556	0.32	0.47	0.00	1.00
Admitted in 1996	1,630	0.33	0.47	0.00	1.00
Admitted in 1997	1,572	0.32	0.47	0.00	1.00
Admitted in 1998 or Later	122	0.03	0.16	0.00	1.00
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	2,936	0.60	0.49	0.00	1.00
Age: 25-30 Years Old	1,577	0.32	0.47	0.00	1.00
Age: Older than 30 Years Old	367	0.08	0.26	0.00	1.00
Race					
White	3,945	0.81	0.39	0.00	1.00
Black/Mulato	735	0.15	0.36	0.00	1.00
Asian	153	0.03	0.17	0.00	1.00
Native	44	0.01	0.09	0.00	1.00
SES Related Variables					
Family Income					
Income: <= R$ 600	233	0.05	0.21	0.00	1.00
Income: R$ 601 to R$ 2.000	1,458	0.30	0.46	0.00	1.00
Income: R$ 2.001 to R$ 4.000	1,520	0.31	0.46	0.00	1.00
Income: R$ 4.001 to R$ 10.000	1,300	0.27	0.44	0.00	1.00
Income: >= R$ 10.001	369	0.08	0.26	0.00	1.00
Parental Education					
Father Having HS Degree	1,167	0.24	0.43	0.00	1.00
Father Having College Degree	2,212	0.45	0.50	0.00	1.00
Mother Having HS Degree	1,630	0.33	0.47	0.00	1.00
Mother Having College Degree	1,748	0.36	0.48	0.00	1.00
Type of High Scholl Attended					
All Private HS	2,737	0.56	0.50	0.00	1.00
All Public HS	1,520	0.31	0.46	0.00	1.00
Some Public / Some Private	623	0.13	0.33	0.00	1.00
Involvement in College Activities					
Student Involvement		4.29	1.26	0.00	6.00

Appendix B.1.4 -Mean, Standard Deviation, Minimum, Maximum of Student-Level Variables - Program: Dentistry

Variable	N	Mean	SD	Min.	Max.
Student-Level Variables (N=6595)					
Student Test Score (ENC)					
Test Score (Multiple-Choice)		55.78	10.13	0.00	90.00
Test Score (Open-Question Format)		52.28	12.73	0.00	95.00
Test Score (All Questions)		54.03	9.25	17.50	87.50
Standardized Test Score		500.00	100.00	104.87	861.96
Admission Cohort					
Admitted in 1995 or Before	178	0.03	0.16	0.00	1.00
Admitted in 1996	703	0.11	0.31	0.00	1.00
Admitted in 1997	3,787	0.57	0.49	0.00	1.00
Admitted in 1998 or Later	1,927	0.29	0.45	0.00	1.00
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	5,376	0.82	0.39	0.00	1.00
Age: 25-30 Years Old	1,099	0.17	0.37	0.00	1.00
Age: Older than 30 Years Old	120	0.02	0.13	0.00	1.00
Race					
White	5,590	0.85	0.36	0.00	1.00
Black/Mulato	621	0.09	0.29	0.00	1.00
Asian	355	0.05	0.23	0.00	1.00
Native	27	0.00	0.06	0.00	1.00
SES Related Variables					
Family Income					
Income: <= R$ 600	152	0.02	0.15	0.00	1.00
Income: R$ 601 to R$ 2.000	1,222	0.19	0.39	0.00	1.00
Income: R$ 2.001 to R$ 4.000	2,388	0.36	0.48	0.00	1.00
Income: R$ 4.001 to R$ 10.000	2,295	0.35	0.48	0.00	1.00
Income: >= R$ 10.001	538	0.08	0.27	0.00	1.00
Parental Education					
Father Having HS Degree	1,515	0.23	0.42	0.00	1.00
Father Having College Degree	3,664	0.56	0.50	0.00	1.00
Mother Having HS Degree	2,119	0.32	0.47	0.00	1.00
Mother Having College Degree	3,116	0.47	0.50	0.00	1.00
Type of High Scholl Attended					
All Private HS	4,734	0.72	0.45	0.00	1.00
All Public HS	981	0.15	0.36	0.00	1.00
Some Public / Some Private	880	0.13	0.34	0.00	1.00
Involvement in College Activities					
Student Involvement		4.84	1.07	0.00	6.00

Appendix B.2.1 - Number of Programs, Number of Students, Mean and Standard Deviation of Standardized Test Scores (ENC) by Program Evaluation Ratings (ACO) - Program: Administration

Results of ACO Evaluations	Number of Programs		Number of Students		Standardized Test Score		Selectivity	
					Mean	SD	Mean	SD
All programs	327		30,776		498.76	50.80	4.14	3.93
Faculty Evaluation from ACO					***		***	
Insufficient	19	6%	715	2%	491.76	44.44	2.76	2.67
Regular	166	51%	13,033	42%	484.83	40.55	3.04	2.12
Good	122	37%	14,658	48%	508.88	53.25	4.96	4.53
Very Good	20	6%	2,370	8%	559.38	61.58	9.60	6.37
Curriculum Evaluation from ACO					***		**	
Insufficient	56	17%	3,471	11%	477.90	42.99	3.42	3.03
Regular	111	34%	10,052	33%	490.21	37.07	3.44	3.24
Good	112	34%	12,292	40%	503.71	54.19	4.53	4.17
Very Good	48	15%	4,961	16%	531.36	61.22	5.69	5.12
Facility Evaluation from ACO					*			
Insufficient	28	9%	1,237	4%	487.11	43.07	4.00	3.78
Regular	75	23%	4,819	16%	489.02	48.97	4.59	4.20
Good	129	39%	12,128	39%	499.14	47.33	3.97	3.49
Very Good	95	29%	12,592	41%	509.38	56.98	4.06	4.34

* p<.05 ** p<.01 *** p<.001

Appendix B.2.2 - Number of Programs, Number of Students, Mean and Standard Deviation of Standardized Test Scores (ENC) by Program Evaluation Ratings (ACO) - Program: Law

Results of ACO Evaluations	Number of Programs		Number of Students		Standardized Test Score		Selectivity	
					Mean	SD	Mean	SD
All programs	189		34,834		507.71	46.94	10.48	8.50
Faculty Evaluation from ACO					***		***	
Insufficient	52	28%	8,124	23%	500.00	34.37	8.33	6.12
Regular	63	33%	10,680	31%	500.99	48.53	9.42	7.23
Good	59	31%	12,794	37%	509.07	48.14	10.37	7.36
Very Good	15	8%	3,236	9%	557.27	46.81	22.81	13.87
Curriculum Evaluation from ACO								
Insufficient	60	32%	9,667	28%	506.15	47.78	11.30	8.89
Regular	47	25%	8,278	24%	503.94	40.34	8.27	5.34
Good	45	24%	9,893	28%	508.06	51.70	11.10	8.41
Very Good	37	20%	6,996	20%	514.60	48.47	11.20	10.78
Facility Evaluation from ACO							*	
Insufficient	47	25%	5,825	17%	515.68	48.35	13.34	8.98
Regular	26	14%	3,392	10%	494.84	41.49	8.58	7.75
Good	77	41%	18,219	52%	504.06	43.48	8.91	6.89
Very Good	39	21%	7,398	21%	513.90	53.86	11.39	10.34

* p<.05 ** p<.01 *** p<.001

Appendix B.2.3 - Number of Programs, Number of Students, Mean and Standard Deviation of Standardized Test Scores (ENC) by Program Evaluation Ratings (ACO) - Program: Civil Engineering

Results of ACO Evaluations	Number of Programs	Number of Students	Standardized Test Score		Selectivity	
			Mean	SD	Mean	SD
All programs	102	4,880	499.10	59.32	4.63	3.92
Faculty Evaluation from ACO			***		***	
Insufficient	15 15%	607 12%	464.41	44.30	3.43	3.09
Regular	40 39%	1,489 31%	468.30	54.88	2.20	2.06
Good	31 30%	1,819 37%	514.70	29.41	6.41	3.28
Very Good	16 16%	965 20%	578.39	36.84	8.40	4.79
Curriculum Evaluation from ACO			***		**	
Regular	28 27%	1,348 28%	472.63	54.20	3.33	2.89
Good	50 49%	2,311 47%	492.82	49.07	4.22	3.39
Very Good	24 24%	1,221 25%	543.06	62.59	7.00	4.99
Facility Evaluation from ACO			***			
Regular	26 25%	1,223 25%	490.49	59.80	4.44	3.08
Good	54 53%	2,429 50%	484.01	48.47	4.02	3.55
Very Good	22 22%	1,228 25%	546.30	61.04	6.38	5.16

* p<.05 ** p<.01 *** p<.001

Appendix B.2.4 - Number of Programs, Number of Students, Mean and Standard Deviation of Standardized Test Scores (ENC) by Program Evaluation Ratings (ACO) - Program: Dentistry

Results of ACO Evaluations	Number of Programs		Number of Students		Standardized Test Score		Selectivity	
					Mean	SD	Mean	SD
All programs	80		6,595		507.52	47.36	17.46	10.70
Faculty Evaluation from ACO					*		**	
Regular	17	21%	1,396	21%	486.85	33.36	11.44	6.84
Good	41	51%	3,300	50%	503.97	51.44	16.83	10.28
Very Good	22	28%	1,899	29%	530.11	40.54	23.28	11.31
Curriculum Evaluation from ACO					*			
Insufficient	12	15%	941	14%	471.62	47.15	14.02	6.61
Regular	24	30%	1,911	29%	515.21	46.23	17.63	12.56
Good	27	34%	2,142	32%	516.76	44.90	16.67	10.11
Very Good	17	21%	1,601	24%	507.35	44.69	20.90	10.95
Facility Evaluation from ACO					*			
Insufficient	11	14%	814	12%	467.19	33.20	15.10	6.60
Regular	34	43%	2,559	39%	513.47	39.41	15.34	10.76
Good	21	26%	1,899	29%	521.86	40.62	20.36	12.33
Very Good	14	18%	1,323	20%	503.29	66.92	20.10	9.84

* p<.05 ** p<.01 *** p<.001

Appendix B.3.1 - Mean of Student-Level Variables, by Faculty Evaluation Ratings - Program: Administration

Variables	Faculty Evaluation Ratings				Total
	"Insuffi-cient"	"Regular"	"Good"	"Very Good"	
Admission Cohort					
Admitted in 1995 or Before	0.08	0.10	0.12	0.13	0.11
Admitted in 1996	0.08	0.11	0.12	0.15	0.12
Admitted in 1997	0.20	0.30	0.30	0.38	0.31
Admitted in 1998 or Later	0.64	0.49	0.45	0.34	0.46
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	0.48	0.49	0.54	0.63	0.53
Age: 25-30 Years Old	0.27	0.30	0.29	0.25	0.29
Age: Older than 30 Years Old	0.25	0.20	0.17	0.12	0.18
Race					
White	0.79	0.83	0.83	0.82	0.83
Black/Mulato	0.18	0.14	0.12	0.13	0.13
Asian	0.01	0.02	0.04	0.04	0.03
Native	0.02	0.01	0.01	0.01	0.01
SES Related Variables					
Family Income					
Income: <= R$ 600	0.10	0.05	0.03	0.03	0.04
Income: R$ 601 to R$ 2.000	0.46	0.36	0.28	0.23	0.32
Income: R$ 2.001 to R$ 4.000	0.28	0.35	0.34	0.31	0.34
Income: R$ 4.001 to R$ 10.000	0.13	0.20	0.26	0.29	0.23
Income: >= R$ 10.001	0.03	0.05	0.08	0.15	0.07
Father Educational Level					
Father Having less than HS Degree	0.66	0.56	0.46	0.32	0.49
Father Having HS Degree	0.22	0.23	0.23	0.24	0.23
Father Having College Degree	0.12	0.22	0.31	0.45	0.28
Type of High Scholl Attended					
All Private HS	0.21	0.37	0.47	0.58	0.43
All Public HS	0.63	0.46	0.38	0.31	0.41
Some Public / Some Private	0.16	0.16	0.15	0.11	0.15
Involvement in College Activities					
Student Involvement	4.20	3.66	3.66	3.91	3.69

Appendix B.3.2 - Mean of Student-Level Variables, by Faculty Evaluation Ratings - Program: Law

Variables	Faculty Evaluation Ratings				Total
	"Insuffi-cient"	"Regular"	"Good"	"Very Good"	
Admission Cohort					
Admitted in 1995 or Before	0.08	0.07	0.10	0.10	0.08
Admitted in 1996	0.20	0.21	0.27	0.22	0.23
Admitted in 1997	0.71	0.69	0.61	0.65	0.66
Admitted in 1998 or Later	0.02	0.03	0.02	0.04	0.02
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	0.53	0.54	0.57	0.66	0.56
Age: 25-30 Years Old	0.25	0.24	0.24	0.19	0.24
Age: Older than 30 Years Old	0.22	0.23	0.19	0.15	0.20
Race					
White	0.84	0.82	0.86	0.83	0.84
Black/Mulato	0.13	0.15	0.11	0.13	0.13
Asian	0.01	0.02	0.02	0.02	0.02
Native	0.01	0.01	0.01	0.01	0.01
SES Related Variables					
Family Income					
Income: <= R$ 600	0.05	0.04	0.04	0.03	0.04
Income: R$ 601 to R$ 2.000	0.31	0.30	0.26	0.19	0.28
Income: R$ 2.001 to R$ 4.000	0.33	0.33	0.32	0.28	0.32
Income: R$ 4.001 to R$ 10.000	0.25	0.26	0.29	0.35	0.28
Income: >= R$ 10.001	0.06	0.06	0.09	0.15	0.08
Father Educational Level					
Father Having less than HS Degree	0.39	0.39	0.36	0.21	0.36
Father Having HS Degree	0.22	0.23	0.22	0.18	0.22
Father Having College Degree	0.40	0.38	0.42	0.61	0.42
Type of High Scholl Attended					
All Private HS	0.51	0.52	0.52	0.69	0.53
All Public HS	0.32	0.32	0.32	0.21	0.31
Some Public / Some Private	0.17	0.16	0.17	0.10	0.16
Involvement in College Activities					
Student Involvement	4.11	4.15	4.00	3.99	4.07

Appendix B.3.3 - Mean of Student-Level Variables, by Faculty Evaluation Ratings - Program: Civil Engineering

Variables	Faculty Evaluation Ratings				Total
	"Insuffi-cient"	"Regular"	"Good"	"Very Good"	
Admission Cohort					
Admitted in 1995 or Before	0.28	0.39	0.30	0.27	0.32
Admitted in 1996	0.34	0.34	0.32	0.35	0.33
Admitted in 1997	0.33	0.24	0.36	0.37	0.32
Admitted in 1998 or Later	0.05	0.03	0.02	0.01	0.03
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	0.48	0.48	0.66	0.74	0.60
Age: 25-30 Years Old	0.38	0.39	0.29	0.24	0.32
Age: Older than 30 Years Old	0.14	0.12	0.04	0.02	0.08
Race					
White	0.71	0.84	0.80	0.85	0.81
Black/Mulato	0.27	0.13	0.16	0.10	0.15
Asian	0.00	0.02	0.04	0.05	0.03
Native	0.02	0.01	0.01	0.00	0.01
SES Related Variables					
Family Income					
Income: <= R$ 600	0.06	0.05	0.05	0.04	0.05
Income: R$ 601 to R$ 2.000	0.41	0.33	0.28	0.22	0.30
Income: R$ 2.001 to R$ 4.000	0.30	0.34	0.31	0.27	0.31
Income: R$ 4.001 to R$ 10.000	0.19	0.23	0.27	0.35	0.27
Income: >= R$ 10.001	0.03	0.05	0.08	0.12	0.08
Father Educational Level					
Father Having less than HS Degree	0.40	0.40	0.26	0.18	0.31
Father Having HS Degree	0.26	0.25	0.26	0.19	0.24
Father Having College Degree	0.34	0.35	0.48	0.63	0.45
Type of High Scholl Attended					
All Private HS	0.46	0.46	0.61	0.68	0.56
All Public HS	0.36	0.38	0.28	0.24	0.31
Some Public / Some Private	0.18	0.16	0.11	0.08	0.13
Involvement in College Activities					
Student Involvement	4.12	4.06	4.41	4.53	4.29

Appendix B.3.4 - Mean of Student-Level Variables, by Faculty Evaluation Ratings - Program: Dentistry

Variables	Faculty Evaluation Ratings				Total
	"Insuffi-cient"	"Regular"	"Good"	"Very Good"	
Admission Cohort					
Admitted in 1995 or Before	0.01	0.02	0.03	0.02	0.03
Admitted in 1996	0.20	0.13	0.11	0.08	0.11
Admitted in 1997	0.53	0.55	0.64	0.47	0.57
Admitted in 1998 or Later	0.26	0.30	0.21	0.42	0.29
Demographic Characteristics					
Age					
Age: Up to 24 Years Old	0.80	0.80	0.80	0.86	0.82
Age: 25-30 Years Old	0.19	0.17	0.18	0.13	0.17
Age: Older than 30 Years Old	0.01	0.02	0.02	0.01	0.02
Race					
White	0.85	0.86	0.85	0.84	0.85
Black/Mulato	0.12	0.11	0.10	0.08	0.09
Asian	0.02	0.02	0.05	0.08	0.05
Native	0.01	0.00	0.01	0.00	0.00
SES Related Variables					
Family Income					
Income: <= R$ 600	-	0.02	0.02	0.02	0.02
Income: R$ 601 to R$ 2.000	0.11	0.19	0.19	0.18	0.19
Income: R$ 2.001 to R$ 4.000	0.33	0.38	0.36	0.35	0.36
Income: R$ 4.001 to R$ 10.000	0.45	0.34	0.34	0.36	0.35
Income: >= R$ 10.001	0.11	0.07	0.08	0.09	0.08
Father Educational Level					
Father Having less than HS Degree	0.17	0.23	0.22	0.20	0.21
Father Having HS Degree	0.23	0.25	0.23	0.21	0.23
Father Having College Degree	0.60	0.52	0.54	0.59	0.56
Type of High Scholl Attended					
All Private HS	0.79	0.70	0.72	0.73	0.72
All Public HS	0.10	0.13	0.16	0.14	0.15
Some Public / Some Private	0.11	0.18	0.12	0.13	0.13
Involvement in College Activities					
Student Involvement	4.30	4.91	4.85	4.82	4.84

Appendix B.3 - Mean Test Score, Initial and Final Beta Coefficients, by Educational Conditions and by Academic Field

Educational Conditions	Administration			Law			Civil Engineering			Dentistry		
	Mean Test Score	Beta (Initial)	Beta (Final)	Mean Test Score	Beta (Initial)	Beta (Final)	Mean Test Score	Beta (Initial)	Beta (Final)	Mean Test Score	Beta (Initial)	Beta (Final)
Faculty												
Insufficient	491	-69.28 ***	-1.25	496	-49.20 ***	-5.68 *	446	-127.91 ***	-61.80 ***	[a]	[a]	[a]
Regular	486	-74.53 ***	-17.61 ***	497	-47.68 ***	-11.88 ***	461	-113.50 ***	-40.44 ***	484	-38.09 ***	-15.45 ***
Good	503	-57.46 ***	-15.45 ***	493	-51.70 ***	-13.67 ***	511	-63.76 ***	-31.58 ***	495	-27.11 ***	-15.94 ***
Very Good	561	[b]	[b]	545	[b]	[b]	574	[b]	[b]	522	[b]	[b]
Curriculum												
Insufficient	478	-55.76 ***	-20.45 ***	497	-4.88 **	-3.34	[a]	[a]	[a]	470	-24.80 ***	2.36
Regular	498	-35.45 ***	-12.15 ***	503	1.03	0.45	455	-93.05 ***	-29.66 ***	506	11.03 **	14.75 ***
Good	495	-38.57 ***	-11.07 ***	499	-2.92	-1.97	500	-47.90 ***	-14.23 ***	512	16.90 ***	13.66 ***
Very Good	533	[b]	[b]	502	[b]	[b]	548	[b]	[b]	495	[b]	[b]
Facility												
Insufficient	480	-24.95 ***	8.98 **	516	7.11 ***	5.16 **	[a]	[a]	[a]	461	-31.60 ***	-0.30
Regular	496	-8.70 ***	-1.77	489	-19.44 ***	-11.40 ***	476	-64.74 ***	2.79	505	12.59 ***	24.38 ***
Good	499	-5.68 ***	2.19	494	-14.74 ***	3.31 *	491	-50.19 ***	-5.25	514	21.31 ***	16.51 ***
Very Good	505	[b]	[b]	508	[b]	[b]	541	[b]	[b]	493	[b]	[b]

[a] Due to low number of cases, programs with insufficient evaluation score were considered as with regular conditions

[b] Used as reference group in the regressions

REFERENCES

Alexander, J. M., & Stark, J. S. (1986). *Focusing on student academic outcomes: A working paper.* National Center for Research to Improve Postsecondary Teaching and Learning. Ann Arbor, Michigan: NCRIPTAL.

Anderson, K. (1984). *Institutional differences in college effects.* Unpublished paper, Florida Atlantic University, Boca Raton.

Astin, A. W. (1970). The methodology of research on college impact, part one. *Sociology of Education, 43*(3), 223-254.

Astin, A. W. (1974). Measuring the outcomes of higher education. In H. R. Bowen (Ed.), *Evaluating institutions for accountability* (New Directions for Institutional Research, No. 1). San Francisco: Jossey-Bass.

Astin, A. W. (1977). *Four critical years: Effects of college on beliefs, attitudes, and knowledge.* San Francisco: Jossey-Bass.

Astin, A. W. (1982). *Minorities in American higher education.* San Francisco: Jossey-Bass.

Astin, A. W. (1985). *Achieving educational excellence.* San Francisco: Jossey-Bass.

Astin, A. W. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education.* New York: Maxwell Macmillan.

Astin, A. W. (1993). *What matters in college? Four critical years revisited.* San Francisco: Jossey-Bass.

Astin, A. W. (1999). Assessment, student development, and public policy. In S. J. Messick (Ed.), *Assessment in higher education: Issues of access, quality, student development, and public policy.* Mahwah, NJ: Lawrence Erlbaum Associates.

Astin, A. W., & Lee, J. J. (2003). How risky are one-shot cross-sectional assessments of undergraduate students? *Research in Higher Education, 44*(6), 657-672.

Barnett, R. (1994). The idea of quality: Voicing the educational. In G. D. Doherty (Ed.), *Developing quality systems in higher education.* London: Routledge.

Bastos, L. R., Michael, W. B., & Oliveira, D. T. (2001). A critique of the external product evaluation of the Brazilian system of undergraduate education. *International Review of Education, 47*(1), 151–158.

Billing, D., & Thomas, H. (2000). The international transferability of quality assessment systems for higher education: the Turkish experience. *Quality in Higher Education, 6*(1), 31-40.

Bishop, J. H., Manã, F., & Bishop, M. (2001). How external exit exams spur achievement. *Educational Leadership, 59*(1), 58-63.

Bogue, E. G. (1998). Quality assurance in higher education: The evolution of systems and design ideals. In G. Gaither (Ed.), *Quality assurance in higher education* (New Directions for Institutional Research, No. 99, pp. 7-18). San Francisco: Jossey-Bass.

Bogue, E. G., & Saunders, R. L. (1992). *The evidence for quality: Strengthening the tests of academic and administrative effectiveness*. San Francisco: Jossey-Bass.

Bond, L. (1981). Bias in mental tests. In L. Bond (Ed.), *Issues in testing: Coaching, disclosure, and ethnic bias (*New directions for testing and measurement series, No. 11, pp. 55-77). San Francisco: Jossey-Bass.

Bosker, R. J., & Scheerens, J. (1994). Alternative models of school effectiveness put to the test. *International Journal of Educational Research, 21*, 159–180.

Bourdieu, P. (1973). Cultural reproduction and social reproduction. In R. Brown (ed.). *Knowledge, education and cultural change* (pp. 487-510). London: Tavistock.

Braxton, John M. (1996). Contrasting perspectives on the relationship between teaching and research. In John M. Braxton (Ed.) *Faculty teaching and research: Is there a conflict?* (New Directions for Institutional Research, no. 90, pp 5-14). San Francisco: Jossey-Bass.

Brennan, J., & Shah, T. (1997). Quality assessment, decision-making and institutional change. *Tertiary education and management, 3*(2), 157-164.

Brown, R., & DeCoster, D. (1982). *Mentoring-transcript systems for promoting student growth*. San Francisco: Jossey-Bass.

Burke, J. C., & Minassians, H. P. (2002). The new accountability: From regulation to results. In J. C. Burke, & H. P. Minassians (Eds.), *Reporting higher education results: Missing links in the performance chain.* (New Directions for Institutional Research, No. 116, pp. 5-19). San Francisco: Jossey-Bass.

Caldeira, E., Kraemer, M. E. P., & Vasconcelos, C. B. (2000). Constituindo a avaliação processual no curso de economia pelo exame nacional de cursos. *Revista Brasileira de Estudos Pedagógicos, 81*(198), 347-365.

Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, California: Sage.

Chernay, G. (1990). *Accreditation and the role of council on postsecondary accreditation (COPA)*. Washington, D.C.: COPA.

Chickering, A, W., & Gamson, Z. F. (1991). Seven principles for good practice in undergraduate education. In A. W. Chickering & Z. F. Gamson (Eds), *Applying the seven principles of good practice in undergraduate education* (New Directions for Teaching and Learning, 47, pp. 63-69). San Francisco: Jossey-Bass.

Clark, B., Heist, P., McConnell, M., Trow, M., & Yong, G. (1972). *Students and colleges: Interaction and change*. Berkeley: University of California, Center for Research and Development in Higher Education.

Clewell, B. C., & Ficklen, M. (1986). *Improving minority retention in higher education: A search for effective institutional practices*. Princeton, NJ: Educational Testing Service.

Colbeck, C. (1998). Merging in a seamless blend: How faculty integrate teaching and research. *The Journal of Higher Education 69*(6), 647-671.

Conrad, C. F., & Pratt, A. M. (1985). Designing quality. *Journal of higher education, 56*(6), 601-622.

Cope, R. G., Hannah, W. (1975). *Revolving college doors: The causes and consequences of dropping out, stopping out, and transferring*. New York: John Wiley & Sons.

Creemers, B. P. M., & Scheerens, J. (1994). Developments in the educational effectiveness research programme. *International Journal of Educational Research, 21*, 125–140.

Cross, K. P. (1991). College teaching: What do we know about it? *Innovative Higher Education, 16*, 7-25.

De Graff, N. D., & De Graff, P. M., & Kraaykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: A refinement of cultural capital perspective. *Sociology of Education, 73*(2), 92-111.

Dey, E. L, Hurtado, S., Rhee, B.S., Inkelas, K. K., Wimsatt, L. A., & Guan, F. (1997). *Improving research on postsecondary student outcomes: A review of the strengths and limitations of national data resources*. National Center for Postsecondary Improvement, Stanford University.

Dias Sobrinho, J. (1999). Exames gerais, provão e avaliação educativa. *Avaliação, 4*(13), 27-50.

DiMaggio, P. (1982). Cultural capital and school success: the impact of status culture participation on the grades of U.S. High School Students. *American Sociological Review, 47*(2), 189-201.

Dumais, S. A. (2002). Cultural capital, gender, and school success: The role of habitus. *Sociology of education, 75*(1), 44-68

Dundar, H., & Lewis, D. (1995). Departmental productivity in American universities: Economies of scale and scope. *Economics of Education Review 14*, 119-144.

Durham, E. R. (1993). Uma política para o ensino superior. *Documentos de Trabalho Nupes, 2/93*. Sao Paulo: Núcleo de Pesquisas sobre Ensino Superior da Universidade de São Paulo.

Eitle, T. M., & Eitle, D. J. (2002). Race, cultural capital, and the educational effects of participation in sports. *Sociology of education. 75*(2), 123-146.

El-Khawas, E. (2002). Quality assurance for higher education: shaping effective policy in developing countries. In D. W. Chapman & A. E. Austin (Eds.), *Higher education in the developing world: changing contexts and institutional responses* (pp. 197-216). London: Greenwood Press.

Ewell, P. (1994). Developing statewide performance indicators for higher education: Policy themes and variations. In S. S. Ruppert (Ed.), *Charting higher education accountability: a source book on state-level performance indicators*. Denver: Education Commission of the States.

Ewell, P. (2001). Statewide testing in higher education. *Change, 33*(2), 20-27.

Feldman, K. (1987). Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: A review and exploration. *Research in Higher Education 26*(3), 227-98.

Ferrer, W. M. H. (2001). *O exame nacional de cursos sob a ótica da grande imprensa paulista.* Unpublished doctoral thesis, Universidade Estadual Paulista "Júlio de Mesquita Filho".

Fetters, W. B. (1977). *Withdrawal from institutions of higher education: An appraisal with longitudinal data involving diverse institutions*. Washington, DC: U.S. Government Printing Office.

Flaugher, R. (1974). *Bias in testing: A review and discussion.* TM report no. 36, Educational Resources Information Center Clearinghouse on Testing Service.

Frazer, M. (1992). Quality assurance in higher education. In A. Craft (Ed.), *Quality assessment in higher education: Proceedings of an international conference in Hong Kong, 1991* (pp. 9-28). London: The Falmer Press.

Freeman, k. (1997). Increasing African Americans' participation in higher education: African American high-school students' perspectives. *The Journal of Higher Education, 68*(5), 523-550.

Freitas, A. (2000). *Exame nacional de cursos: Um instrumento de politica publica. Estudo comparativo do impacto do exame em instituicoes do Distrito Federal.* Unpublished master thesis, Universidade de Brasilia.

Gaither, G. (1998). Editor's notes. In G. Gaither (Ed.), *Quality assurance in higher education* (New Directions for Institutional Research, No. 99, pp. 1-5). San Francisco: Jossey-Bass.

Gates, S. M., Augustine, C. H., Benjamin, R., Bikson, T. K., Kaganoff, T., Levy, D. G., Moini, J. S., & Zimmer, R. W. (2002). *Ensuring quality and productivity in higher education: An analysis of assessment practices.* ASHE-ERIC/Higher Education Report Volume 29, No. 1.

Gomes, A. (2002). Higher education assessment policy in Brazil: Control and massification. *Educação & Sociedade, 23*(80), 275-298.

Harvey, L. (1998). An assessment of past and future approaches to quality in higher education. *Australian Journal of Education, 42*(3), 237-255.

Harvey, L., & Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education, 18*(1), 9-26.

Hearn, J. (1991). Academic and non-academic influences on the college destinations of 1980 high school graduates. *Sociology of Education, 64*(3), 158-71.

Hearn, J. C. (1984). The relative roles of academic, ascribed, and socioeconomic characteristics in college destinations. *Sociology of Education, 57*(1), 22-30.

House, J. D. (1999). *The Effects of Entering Characteristics and college experiences on student satisfaction and degree completion: An application of the input-environment-outcome assessment model.* Presented at the Annual Meeting of the Association for Institutional Research, Minneapolis, NM. (ERIC Document Reproduction Service No. ED429 491)

Hu, S., & Kuh, G. D. (2003). Maximizing what students get out of college: Testing a learning productivity model. *Journal of College Student Development, 44*(2), 185-203.

Instituto Nacional de Estudos e Pesquisas Educacionais. (2001). *Exame Nacional de Cursos: Relatório-síntese 2001*. Brasília: O Instituto.

Jackson, R. (1980). The scholastic aptitude test: A response to Slack and Porter's 'critical appraisal'. *Harvard Education Review, 50*, 382-391.

Jacobi, M., Astin, A., & Ayala, F. (1987). *College student outcomes assessment: A talent development perspective.* (ASHE-ERIC Higher Education Report No. 7). Washington, DC: Association for the Study of Higher Education.

Karabel, J., & Astin, A. (1975). Social class, academic ability, and college quality. *Social Forces, 53*, 381-398.

Kim, M. M., Rhoades, G., & Woodard, Jr., D. B. (2003). Sponsored research versus graduating students? Intervening variables and unanticipated findings in public research universities. *Research in Higher Education, 44*(1), 51-81.

Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.

Koretz, D. M., & Barrom, S. L. (1998). *The validity of gains in scores on the Kentucky instructional results information system (KIRIS)*. Santa Monica, CA: Rand.

Lee, J. J. (2002). Religion and college attendance: Change among students. *The Review of Higher Education,* 25(4), 369-384.

Lenning, O. T., Lee, Y. S., Micek, S. S., & Service, L. A. (1977). *A structure for the outcomes of postsecondary education.* Boulder, CO: National Center for Higher Education Management Systems.

Leslie, L. L., Rhoades, G., & Oaxaca, R. (1999). *Effects of changing revenue patterns on public research universities*. A report to the National Science Foundation, Grant Number 9628325.

Lim, D. (2001). *Quality assurance in higher education: A study of developing countries*. Aldershot: Ashgate.

Lipsey, M. (1987). *Theory as method: Small theories of treatments*. Paper presented at the National Center for Health Services Research Conference, "Strengthening causal interpretations of non-experimental data," Tucson, AZ.

Lively, K. (1992, p. A29). Accountability of colleges and gets renewed scrutiny from state officials. *Chronicle of Higher Education, 39*(2), A21.

Maia Filho, T., Pilati, O., & Lira, S. C. (1998). O Exame nacional de cursos. *Revista Brasileira de Estudos Pedagógicos, 79*(192), 74-91.

Martins, M. (2000). *Impactos do exame nacional de cursos e da qualificação do corpo docente sobre o curso de odontologia da PUC minas.*. Paper presented at the 23ª Annual meeting of ANPED, Caxambu, MG [Online]. Available: http://www.anped.org.br/1115t.doc [2003, March 3].

Martuza, V. R. (1977). *Applying norm-referenced and criterion-referenced measurement in education*. Boston: Allyn and Bacon.

McClelland, K. (1990). Cumulative disadvantage among the highly ambitious. *Sociology of Education, 63*(2), 102-121.

McDonough, P. M. (1994). Buying and selling higher education: The social construction of the college applicant. *The Journal of Higher Education, 65*(4), 427-446.

Mehrens, W. A., & Lehmann, I. J. (1987). *Using standardized tests in education* (4th Ed.). New York: Longman

Merton, R. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago Press.

Messick, S. (1981). The controversy over coaching: Issues of effectiveness and equity. In L. Bond (Ed.), *Issues in testing: Coaching, disclosure, and ethnic bias (*New directions for testing and measurement series, No. 11, pp. 21-53). San Francisco: Jossey-Bass.

Milem, J. M., Berger, J. B., & Dey, E. L. (2000). Faculty time allocation: A study of change over twenty years. *The Journal of Higher Education, 71*(4), 454-475.

Ministério da Educação e do Desporto. (1999). *Avaliação das condições de ofertas de cursos de graduação: Relatório síntese 1999*. Brasília: MEC, SESu.

Neave, G., & van Vught, F. (1991). *Prometheus bound: The changing relationship between government and higher education in Western Europe*. Oxford: Pergamon.

Nogueira, C. M. M., & Nogueira, M. A. P. (2002). Bourdieu's sociology of education: Limits and contributions. *Educação & Sociedade, 23*(78), 15-35.

Nordvall, R. C., & Braxton, J. M. (1996). An alternative definition of quality of undergraduate college education: Toward usable knowledge for improvement. *Journal of Higher Education, 67*(5), 483-497.

Osborne, Jason W. (2000). Advantages of Hierarchical Linear Modeling. *Practical Assessment, Research & Evaluation, 7*(1). Available on-line at http://search.epnet.com/direct.asp?an=EJ638489&db=eric&site=ehost.

Pascarella, E. T. (1985). College environmental influences on learning and cognitive development: A critical review and synthesis. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research, Vol. 1*. New York: Agathon.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students*. San Francisco: Jossey-Bass.

Patrick, W. J., & Stanley, E. C. (1998). Teaching and research quality indicators and the shaping of higher education. *Research in Higher Education, 39*(1), 19-41.

Perricci, C. (1980). Gender and achievement: The early careers of college graduates. *Sociological Focus, 13*, 99-111.

Persell, C. H., Catsambis, S., & Cookson, P. W. Jr. (1992). Differential asset conversion: Class and gendered pathways to selective colleges. *Sociology of Education, 65*(3), 208-225.

Popham, W. (1999). Why standardized tests don't measure educational quality. *Educational Leadership, 56*(6), 8-15.

Popham, W. (2001). Teaching to the test? *Educational Leadership, 58*(6), 16-20.

Rear, J. (1994). Institutional responses in British higher education. In Westerheijden, D. F., J. Bennam, & P. A. M. Maassen (Eds.), *Changing the contexts of quality assessment: Recent developments in Western European higher education*. Utrecht: Lemma.

Rhoades, G. (2000). The changing role of faculty. In Joseph Losco and Brian L. Fife (eds.), *Higher education in transition: The challenges of the new millennium*. Westport, Connecticut: Bergin and Garvey.

Rhoades, G., & Sporn, B. (2002). Quality assurance in Europe and the U.S.: Professional and political economic framing of higher education policy. *Higher education, 43*, 355-390.

Richards, C. E. (1988). Indicators of three types of educational monitoring systems: Implications for design. *Phi Delta Kappan, 69*(2), 495-499.

Ristoff, D. I. (1996). Princípios do programa de avaliação institutional. *Avaliação, 1*(1), 47-53.

Sampaio, H. (1998). Estabelecimentos de ensino superior privados: A heterogeneodade e qualidade. Sao Paulo: *Documentos de Trabalho Nupes, 8*/98. Nucleo de Pesquisas sobre Ensino Superior da Universidade de Sao Paulo.

Santos Filho, J. C. (1999). Análise teórico-política do exame nacional de cursos. *Avaliação, 4*(13), 9-24.

Santos, W. (2001). *O provao: Para alem do discurso official - Representacao do professor e anatomia do modelo.* Unpublished doctoral thesis, Pontificia Universidade Catolica de Sao Paulo.

Sax, L. J. (2001). Undergraduate science majors: Gender differences in who goes to graduate school. *The Review of Higher Education,* 24(2), 153–172.

Schmoker, M. (2000). The results we want. *Educational Leadership 57*(5), 62-65.

Segenreich, S. C. D. (2000). *O significado da atual avaliação externa do MEC para as universidades brasileiras* [Online]. Available: http://www.calpe.ced.ufsc.br/download/provao/segenreich.zip [2003, March 3].

Seneca, J. J., & Taussig, M. K. (1987). Educational quality, access, and tuition policy at state universities. *Journal of Higher Education, 58*(1), 25-37.

Shepard, L. A. (1981). Identifying bias in test items. In L. Bond (Ed.), *Issues in testing: Coaching, disclosure, and ethnic bias (*New directions for testing and measurement series, No. 11, pp. 79-104). San Francisco: Jossey-Bass.

Shwartzman, S. (2003). *Quality, standards and globalization in higher education* [Online]. Available: http://www.schwartzman.org.br/simon/standards.pdf [2003, March 3]

Silva, J. C. (2001). *O exame nacional de cursos: A utilização de resultados por instituições do Distrito Federal.* Unpublished doctoral thesis, Universidade de Brasilia.

Slack, W. V., & Porter, D. (1980). Training, validity, and the issue of aptitude: A reply to Jackson. *Harvard Education Review, 50,* 392-401.

Soares, J. F., Ribeiro, L. M., & Castro, C. M. (2001). Valor agregado de institutições de ensino superior em Minas Gerais para os cursos de direito, administração e engenharia civil. *Dados, 44*(2). [online]. Available: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0011-52582001000200005&lng=en&nrm=iso.

Stanley, E. C, & Patrick, W. J. (1998). Quality assurance in American and British higher education: A comparison. In G. Gaither (Ed.), *Quality assurance in higher education* (New Directions for Institutional Research, No. 99, pp. 39-56). San Francisco: Jossey-Bass.

Stoecker, J., Pascarella, E., & Wolfe, L. (1988). Persistence in higher education: A nine year test of a theoretical model. *Journal of College Student Development, 29*, 196-209.

Stufflebeam, D. (2001). Evaluation models. In D. Stufflebeam (Ed.), *Evaluation models* (New directions for evaluation, No. 89, pp.7-98). San Francisco: Jossey-Bass.

Tam, M. (2001). Measuring quality and performance in higher education. *Quality in Higher Education, 7(1), 47-54.*

Thomas, S. (1998). *Abordagens de valor agregado para a auto-avaliação escolar no Reino Unido*. Paper presented at the "seminário internacional de avaliação educacional". Brasília: INEP.

Thurmond, V. A., Wambach, K., & Connors, H. R. (2002). Evaluation of student satisfaction: Determining the impact of a web-based environment by controlling for student characteristics. *American Journal of Distance Education, 16*(3), 169-190.

Tinto, V. (1985). Orientation programs. In L. Noel, R. Levitz, & D. Saluri (Eds.), *Increasing student retention: Effective programs and practices for reducing dropout rate*. San Francisco: Jossey-Bass.

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition.* Chicago: University of Chicago Press, 1987

Trow. M. (1984). The analysis of status. In B. R. Clark (Ed.), *Perspectives on higher education: Eight disciplinary and comparative views*. Los Angeles: University of California Press.

Van Vught, F. A. (1994). Intrinsic and extrinsic aspects of quality assessment in higher education. In D. F. Westerheijden,, J. Brennan, & P. A. M. Naassen (Eds.), *Changing contexts of quality assessment in higher education*. Utrecht: LEMMA.

Vargas, H. (2000). *Resultados do provão: Surpresas desconcertantes*. Paper presented at the 23ª Annual meeting of ANPED, Caxambu, MG. [Online]. Available: http://www.anped.org.br/25/posteres/josecarlosrothenp11.rtf [2003, March 3].

Vroeijenstijn, T. (1992). External quality assessment, servant of two masters? The Netherlands university perspective. In A. Craft (Ed.), *Quality assessment in higher education: Proceedings of an international conference in Hong Kong, 1991* (pp. 109-132). London: The Falmer Press.

Vroeijenstijn, T. I. (1994). Preparing for the second cycle: External quality assessment in Dutch universities. In Westerheijden, D. F., J. Bennam, & P. A. M. Maassen (Eds.), *Changing the contexts of quality assessment: Recent developments in Western European higher education*. Utrecht: Lemma.

Walpole, M. (2003). Socioeconomic status and college: How SES affects college experiences and outcomes. *The Review of Higher Education, 27*(1), 45-73.

Westerheijden, D. F., Brennan, J., & Naassen, P. A. M. (1994). *Changing contexts of quality assessment in higher education*. Utrecht: LEMMA.

Willms, J. D. (1992). *Monitoring school performance.* Washington, D.C.: The Falmer Press.

Woodhouse, D. (1995). Efficient quality systems. *Assessment & Evaluation in Higher Education, 20*(1), 15-24.

Young, P. S., & Fischer, N. M. (1996). *Identifying undergraduate and post-college characteristics that may affect alumni giving*. Paper presented at the Annual Forum of the Association for Institutional Research, Albuquerque, NM. (ERIC Document Reproduction Service No. ED397 748)

Zweigenhaft, R. L. (1994). Prep school and public school graduates of Harvard. *Journal of higher education, 64*(2), 211-255.