A GEOGRAPHIC-INFORMATION-SYSTEMS-BASED APPROACH TO ANALYSIS OF CHARACTERISTICS PREDICTING STUDENT PERSISTENCE AND GRADUATION

By

Chris Ousley

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

2010

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Chris A. Ousley entitled A Geographic-Information-Systems-Based Approach to Analysis of Characteristics Predicting Student Persistence and Graduation and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: 4/22/10
Dr. Cecilia Rios Aguilar

_____ Date: 4/22/10
Dr. Paul Kohn

_____ Date: 4/22/10
Dr. Jenny Lee

_____ Date: 4/22/10
Dr. Jeff Milem

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies 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.

_____ Date: 4/22/10
Dissertation Director: Dr. Cecilia Rios Aguilar

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 allowable without special permission, provided that accurate acknowledgement of source is made. Request for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of this material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Chris A. Ousley

ACKNOWLEDGEMENTS

"Take away the dross from the silver, and there shall come forth a vessel for the finer." Proverbs 25:3

Over the course of my education, extraordinary contributions came from extraordinary people, and I thank you for all your help. With a few words, I attempt to express the gratitude of a lifetime.

I happily live in a world of intellectualism, of quantifying the unknown, and testing the questionable, yet in the midst of this, it is an unquestioning faith in God that sustains me. For this privilege, I am most thankful.

To Professor Cecilia Rios-Aguilar, I appreciate your skill and scholarly talent, and I am forever thankful for your kindness, understanding, and patience. Your prompt and articulate feedback turned a lifetime learner into a graduate. The trials of this dissertation were no less, the hill no easier to climb, but at the top, it was nice to have a mentor to gaze out at the vista, to see the rocky path below, and to share the experience of the sun rising on a new day. Thank you.

To Professor Gary Rhoades, who journeyed with me, imparting wisdom until deciding to go save the rest of the world. When it was time for the training wheels to come off, I found the pedaling was not as difficult as I feared. Bumps and bruises awaited, but I had learned to pedal, to steer, and even to do a couple of amazingly awesome tricks. I appreciate all the wisdom you have instilled in me over these many years of friendship.

To professors Jenny Lee, Paul Kohn, Jeff Milem, Sheila Slaughter, John Levin, Doug Woodard, and John Cheslock. At each step of this journey, you individually arose to the occasion, and enabled me to persevere. Thank you for sharing your intelligence, your vision, and your insight into the world of higher education and life.

To my colleague and friend, Yan Xie, thank you for your exceptional attention to detail in assisting me with the data, and for learning alongside me as we boldly dabbled in the discovery of greatness.

To my many other friends and colleagues whose hard earned success has kept me going, I am grateful. Annually, you reminded me of my goal with conference presentations, published articles, and faculty positions obtained. Thank you for your encouragement.

DEDICATION

This dissertation is dedicated to my family. Above all others, to my wife Melissa, who shortly after we married, shared with me a wonderful dream—this is that dream realized. We knew the years would pass regardless of our decision, and so we began our journey. Today, several years have passed and become memories of a life with you by my side. From a fortuitous glance 20 years ago, to this moment, I am madly in love you—the love of my youth, and the wife of all my tomorrows. I have always loved you. Thank you for all you are to me.

To Aiden, who accepted the times when Dad needed to study—who inspires through his outlook on life, his love of nature, and his presence of peacefulness. I hope you're ready for tying shoes, riding bikes, catching bugs, and playing Daddy tag.

To Elliot, who challenges me through his brilliance, who turns the normal on its side to find unique answers, and who captures the light of discovery in a wrinkled up nose, a cute freckled face, and a mischievous smile. How dare you be so adorable when I was trying to write, study, and research!

To my parents, thank you for enduring and encouraging my desire to question convention, and for supporting my dreams over the years. I have never forgotten the feeling of pride I felt in you, as you sacrificed to send me to a seemingly unaffordable university so many years ago. Your selflessness set me off on this journey that has truly changed my life and now that of my family. I know as a parent, it's what we do, but it doesn't make it any less appreciated.

I dedicate this to all of you.

TABLE OF CONTENTS

TABLE OF CONTENTS – Continued

TABLE OF CONTENTS – Continued

TABLE OF CONTENTS – Continued

TABLE OF CONTENTS – Continued

TABLE OF CONTENTS – Continued

TABLE OF CONTENTS – Continued

LIST OF TABLES

LIST OF TABLES – Continued

LIST OF TABLES – Continued

LIST OF FIGURES

ABSTRACT

This study sought to provide empirical evidence regarding the use of spatial analysis in enrollment management to predict persistence and graduation. The research utilized data from the 2000 U.S. Census and applicant records from The University of Arizona to study the spatial distributions of enrollments. Based on the initial results, stepwise logistic regression was used to identify spatially associated student and neighborhood characteristics predicting persistence and graduation.

The findings of this research indicate spatial analysis can be used as a valuable resource for enrollment management. Using a theoretical framework of the forms of capital and social reproduction, cultural and social capital characteristics were found to influence persistence at statistically significant levels. Most notably, the social capital proxy of neighborhood education levels, and the cultural capital proxies of the number of standardized tests a student has taken, and when the application for admission is submitted all significantly influenced a student's probability to persistence and graduate. When disaggregating by race and ethnicity, resident Hispanic students from highly Hispanic neighborhoods were found to persist at higher levels in the first year of college attendance. Also, resident Native Americans were found to have a higher probability to persist when evidencing cultural capital characteristics. Since spatially based student and neighborhood characteristics can be quantified and mapped, target populations can be identified and subsequently recruited, resulting in retention-focused admissions.

CHAPTER ONE: THE PURPOSE OF THIS STUDY

By incorporating spatial analysis into enrollment research, enrollment managers have a new tool for predicting a student's probability to persist and graduate from college. Long before a student begins classes, is offered admission to the institution, or even before a prospect is targeted for recruitment, the institution can identify those prospects most likely to persist by considering the student's neighborhood characteristics. Spatial analysis of the demographic information associated with an individual recruit's home address provides greater context for establishing probabilities related to future student performance. Enrollment managers now can utilize spatial analysis and geodemographic tools to harness the volumes of demographic information from the U.S. Census and admission applications. When merged, these student and neighborhood characteristics provide an additional resource for predicting an individual's probability for persistence and graduation. While selecting only the most probable graduates could improve retention and graduation rates, using this identification also can enable earlier intervention for students most likely to need additional assistance and resources. By using spatial analysis to assess the needs of students before they arrive on campus, the enrollment manager has an additional resource to improve persistence and graduation rates without detrimental effects upon growth, diversity, or academic quality.

While this may sound promising to some, it admittedly may have limited applications for practice. This research explores pre-enrollment characteristics of students, yet several decades of research indicate that a great deal changes in the lives of students after they arrive on a college campus (Pascarella & Terenzini, 2005; Kuh,

Kinzie, Buckley, Bridges, & Hayek, 2006). Once enrolled, each individual uniquely reacts, responds, and interacts with the campus environment (Chickering & Gamson, 1987; Astin, 1991; Kuh, Schuh, Whitt, & Associates, 1991; Chickering & Reisser, 1993). Also, the institutional context provides a unique environment that may, or may not contribute to a student's ability to persist and graduate. These reactions, responses, and interactions are not necessarily measureable prior to enrollment, yet all impact educational outcomes. To examine only those student and neighborhood characteristics existing prior to enrollment implies that post-enrollment events do not impact the student's odds of persisting and graduating. This is not true, as the actions of students post enrollment and the institutional context greatly impact outcomes related to persistence and graduation (Pascarella & Terenzini, 2005; Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006). Because of this, it may appear misguided to attempt to predict outcomes solely on the basis of pre-enrollment characteristics, yet this is the reality in which enrollment managers must operate every day—unsure as to how an individual student will uniquely respond to the college experience and unsure of resources the institution will be able to provide students for the duration of enrollment. It is with an understanding of these limitations, the exclusion of institutional effects upon the student, that this research proceeds to examine pre-enrollment student and neighborhood characteristics that may predict student outcomes. Rather than a point of divergence, it is the intent of this writing, that the study of pre-enrollment characteristics be viewed as a complementary method for informing research and attempting to understand the array of characteristics, behaviors, and actions that influence student persistence and graduation.

The early identification of these student and neighborhood characteristics holds enormous potential for university administrators since low retention rates and low graduation rates pose a major problem for institutions of higher education. From 2004 to 2007, 22% of all students enrolling at public four-year institutions in the U.S. failed to return to the same school for their second enrollment year (NCHEMS, 2009). A national longitudinal study of high school students from 1980 to 1990 found only 47% of students at four-year institutions complete their degree requirements at the institution of original attendance (Berkner, Cuccaro-Alamin, & McCormick, 1996). This statistic has remained at around 45% since 1880 (Tinto, 1982). When considering public institutions alone, the four-year completion rate is 28.1% (Astin & Oseguera, 2003). This public institution rate is critical because while there are more private institutions than public institutions, three-fourths of all students in two- and four-year institutions attend public institutions (National Commission on the Cost of Higher Education, 1998). For the institutions they leave, these students represent the loss of substantial resources invested in their education. Withdrawal also represents lost opportunities such as classroom seats taken by those who failed to graduate and financial resources never allocated to students who were more likely to graduate. Consequently, the institution is forced to spend additional resources—not educating students, but recruiting new students to replace these lost students. Successful retention begins with the successful recruitment of students.

Regarding student recruitment, there exists a great deal of literature derived from post-enrollment studies identifying characteristics of students who persisted. Most scholarly articles that have addressed the recruitment stage focus on enrolling students,

instead of focusing on enrolling those students most likely to persist (Kotler, 1976; Litten, Sullivan, & Brodigan, 1983; Hossler, 2000). A sparse amount of literature addresses the importance of enrolling graduates—students most likely to persist and graduate—as opposed to enrolling students most likely to persist to the second year of enrollment (Hossler, 1984; Hossler & Anderson, 2004; Williams, 1984). Existing literature has explored the importance of numerous student characteristics as they relate to persistence and graduation. These findings include the influence of race/ethnicity, gender, high school GPA, classes taken in high school and scores from standardized tests such as the ACT or SAT (Pascarella & Terenzini, 2005; Astin, Tsui, & Avalos, 1996; Adelman, 2006; Hearn, 1991). Recent work from researchers at The Ohio State University examined the geodemographic characteristics associated with enrollees (Marble, Mora, & Herries, 1995); geodemographic characteristics associated with in-state recruitment opportunities (Marble, Mora, & Granados, 1997); and geodemographic characteristics associated with target markets (Mora, 2003). Still, there is an absence in the literature regarding the use of spatial analysis as it relates to persistence outcomes. To address this gap in the existing enrollment management literature, this study will use spatial analysis to identify pre-enrollment student characteristics related to first year persistence, and will use regression analyses to predict first-year retention, and eventually graduation. In doing so, this research will enable greater use of spatial analysis in recruitment and admissions, which provide the foundation for successful student retention and for improved college graduation rates.

The Problem

Given the fiscally troubled environment of many public college campuses today, enrollment managers have to be cognizant of the return on their investment in new students. Increasingly, there are calls for greater accountability, and these calls come from multiple stakeholders (Zumeta, 2000). Since 1990 when Congress enacted the Student Right-To-Know and Campus Security Act, institutions have been required to report "the completion or graduation rate of certificate- or degree-seeking, full-time, undergraduate students entering such institutions" (U.S. Code, 2007). A September 2006 U.S. Department of Education report from the Commission on the Future of Higher Education notes that a critical shortcoming in American higher education is the failure to graduate students, "at least in part because most colleges and universities don't accept responsibility for making sure that those they admit actually succeed" (USDOE, 2006). To remedy this and other perceived failures, the report calls for greater accountability from providers of American higher education. "There is inadequate transparency and accountability for measuring institutional performance, which is more and more necessary to maintaining public trust in higher education" (USDOE, 2006, p. 14). Retention is viewed as an easily quantifiable measure for assessing institutional performance. Furthermore, improving retention rates is an easy target for accountability proponents because despite continuing research funding and intervention strategies, overall retention rates remain problematic.

Purpose and Significance

Most existing research literature on pre-enrollment factors influencing persistence focuses on the student's background and their match to the institution. Students with proven academic success, goals, and supportive parents tend to persist at higher rates (Bean, 1990). By knowing about the student's background, enrollment managers can make informed decisions on how the student will fare at the institution (Cope, 1981). Past studies have delved into great detail on variables known at the time a student applies to the institution. What they have yet to elucidate are variables that may be gleaned prior to application based on the demographic characteristics of the neighborhood in which an applicant resides. This is where spatial analysis can help. Knowing a student's chances to persist, and identifying those most at risk, can allow enrollment mangers to craft the incoming class more intelligently with respect to retention and graduation. Because spatial analysis utilizes existing census data to establish neighborhood characteristics that later can be associated with a student, it makes the evaluation of a potential student possible months, and even years before the student enrolls.

By utilizing data available from the combination of geographic information science and the U.S. Census, this study seeks to bring a new set of variables into the discussion of factors influencing persistence. In doing so, this dissertation will add variables associated with the applicant's community, that is neighborhood demographics derived from the U.S. Census associated with the geographic location of the applicant. The research will use these data to allow retention outcomes to be considered early in the

recruitment process. Additionally, this study will provide empirical evidence as to the implications for practice.

This research will add data derived from a geographical information system (GIS) to enable early identification of prospective students most likely to persist. This writing establishes a basis for further empirical research into questions about accountability within higher education with respect to an enrollment management office's ability to affect student retention through improved targeting of admissions. This study will address the following research questions:

1. What are the spatial patterns of enrollment for the institution of study with respect to: neighborhood household income levels; neighborhood racial/ethnic compositions; neighborhood education levels; and resident and nonresident enrollees?

2. When using spatial analysis, what are the characteristics associated with a student's probability to persist?

    - How do neighborhood household income levels predict persistence?
    - How do neighborhood racial/ethnic compositions predict persistence?
    - How do neighborhood education levels predict persistence?

3. When using spatial analysis, what are the characteristics associated with a student's probability to graduate?

    - How do neighborhood household income levels predict 4-year graduation rates?

    - How do neighborhood racial/ethnic compositions predict 4-year graduation rates?

    - How do neighborhood education levels predict 4-year graduation rates?

The purpose here is to expand enrollment management literature regarding spatial analysis by providing empirical data. As mentioned earlier, most literature lacks data regarding geographic-related variables known early in the recruitment process (Litten, Sullivan, & Brodigan, 1983). Background variables traditionally have become known to the institution at the time a student becomes a prospective student through the submission of standardized test scores like the ACT or SAT, or they submit an admissions application. As a result, these data are available by virtue of an individual's self-identification as a prospect, rather than the institution finding them. Also, prospect lists from vendors focus on student characteristics indicating those most likely to apply to the institution, but do not focus on the students most likely to persist to graduation. Most admissions literature ends with the matriculation stage and often focuses on access or the recruitment of students in terms of increasing institutional enrollments. (Litten, Sullivan,

& Brodigan, 1983; Hossler, 1999). The purpose of this study is to illuminate discussion on using spatial analysis for retention-focused admissions. Providing enrollment managers with this additional tool for improving student retention may translate into more efficient use of fiscal resources, and provide an early advantage for assisting those students in need of additional campus services. This study allows for the discussion of retention at a much earlier point of consideration for institutions. Instead of retention efforts beginning at the point of orientation as some models posit, consideration of recruiting for retention outcomes starts at the infancy of recruitment, before communication with the prospect takes place.

Organization of the Study

To explore implications of using spatial analysis to better inform student recruitment, this study is organized into six chapters. Chapter Two will discuss literature on spatial analysis and geographical information science, along with their relevance to enrollment management. Then, a review of existing enrollment management literature will be presented with an emphasis on the intersection of enrollment management's subsidiaries, admissions and retention. Chapter Three will provide a discourse on the theoretical frameworks guiding the study. Under the overarching theme of the forms of capital, economic, cultural, and social capital and their relationship to student characteristics will be discussed. Also, the idea that this capital is transferred to the students will be presented through a discussion on social reproduction. Chapter Four presents the research design including methodology, data collection, variable selection, and data analysis. Chapter Five delves into findings from analysis of four years of

quantitative data, representing more than 21,000 cases. Finally, Chapter Six will present the implications of the findings and conclusions of this study. An emphasis will be placed on implications for practice and recommendations for future research.

Summary

The addition of spatial analysis to the enrollment manager's repertoire enables greater understanding of student persistence. By merging geodemographic data associated with a prospective student's neighborhood with applicant data, enrollment managers can better predict those students most likely to persist, and those students most likely in need of additional assistance to persist. At a time with greater calls to improve institutional retention rates, spatial analysis is an additional tool available to better understand student populations, and to improve persistence through a greater understanding of student needs. Using spatial analysis to inform the retention theory in context of enrollment management goals enables the results of this study to be placed in the context of a much larger volume of literature.

CHAPTER TWO: REVIEW OF THE LITERATURE

To understand spatial analysis (used interchangeably with the term geospatial analysis) and its potential for retention-focused enrollment management, several interrelated topics must be presented. I begin this chapter with an overview of spatial analysis literature, including its history and the use of spatial analysis in enrollment management. Next, the topic of enrollment management is presented in greater detail as a unifying theme to understand how the spatial analysis of geodemographic information can inform admissions and retention efforts. By examining selected literature surrounding these topics, this study also explores the concept of institutional fit. I then introduce several student and neighborhood characteristics that relate to persistence and graduation outcomes. In doing so, this chapter provides an understanding as to where these topics may overlap, and informs enrollment management literature regarding the use of spatial analysis to identify those neighborhood demographic characteristics most likely to predict student persistence and graduation.

Spatial Analysis: An Overview

Aspatial data are data that exist independent of a fixed location. In higher education research, this can include student characteristics such as race, GPA, and test scores. For years, researchers have analyzed these aspatial data collected from admission applications and enrollment data to establish probabilities for student persistence (Litten, Sullivan, & Brodigan, 1983; Astin, Tsui, & Avalos, 1996; Astin, 1982; Pascarella & Terenzini, 1980). Conversely, spatial data contain "locational information as well as attribute information" (Fotheringham & Rogerson, 2009, p. 1). Spatial analysis is

different from aspatial analysis because it provides context for the data, and outcomes "are dependent in some way on the locations of the objects being analyzed—if the locations change, the results change" (Fotheringham & Rogerson, 2009; Goodchild & Janelle, 2004, p. 3).

Underlying the concept of spatial analysis is GIS, an acronym that can refer to either geographic information science or a geographic information system (Longley, Goodchild, Maguire, & Rhind, 2005). GIS and spatial analysis are not interchangeable terms. Instead, GIS is the science (or system) that enables a spatial analysis of geographically related data (Martin, 2009). As a system, GIS is comprised of the hardware and software—especially the database containing volumes of data related to a single point on a map (Longley, Goodchild, Maguire, & Rhind, 2005). As a science it refers to the overarching study of geography-based data (Longley, Goodchild, Maguire, & Rhind, 2005). The term GIS is used in both ways within this dissertation depending upon the context.

Spatial Analysis: History

The foundation for modern spatial analysis has been around since Ptolemy (circa 100-178 AD, 1932) expressed physical locations as longitude and latitude based on a spherical earth translated into a flat surface. It would take another 1600 years for mainstream usage within the social sciences (Snow, 1936). Initially, the mapping of trends was used for geographic representations, sometimes incorporating demographic and economic data (Longley, Goodchild, Maguire, & Rhind, 2005). Later, demographic

variables associated with neighborhoods were used by social scientists to examine social and economic change (Wilson, 1987; Massey, Gross, & Shibuya, 1994; Turley, 2003). Also leading the way were medical researchers, who based their work on Hippocrates and his ideas of observing local conditions to understand the context of diseases and their spread (Nuernberger, 2001). One of the most famous examples of this was John Snow's use of a geographical plotting of cholera cases. In 1854, Snow mapped London cholera cases and their proximity to water sources (Snow, 1936). Near a single intersection, an estimated 500 deaths occurred in a few days (Snow, 1936). By mapping the data, the spread of cholera was linked to localized use of contaminated wells, and the epidemic was contained (Snow, 1936). Just as evidence of the epidemic was not randomly distributed, social phenomena "are not randomly distributed across the physical terrain, and as a result, inquiry into the spatial patterning of these phenomena can yield unique insight into their causal dynamics" (Messner & Anselin, 2004, p. 128).

While combining aspatial attributes with spatial attributes began to gain acceptance within the scientific community, it remained impractical for analysis of large data sets due to software and hardware limitations. These limitations relegated GIS to what was primarily a mapping tool used to illustrate geographic features (Berry, 1987). However, technological advances in geographical information system software, data sources, and computer processing now enable larger-scale analysis resolving previous limitations (Longley, Goodchild, Maguire, & Rhind, 2005). GIS as a database manager, enables another software program to draw cartographic representations of data (Peters & MacDonald, 2004). The key skill for managing and manipulating these data is the

linking of spatial and aspatial data to create useful output (Wang, 2006). The industry-leading software for accomplishing this task is ArcGIS, a product of Environmental Systems Research Institute, Inc. (ESRI). This software is used in this study to visualize and analyze spatial data. Once a link is established between spatial and attribute data, GIS allows spatial queries for discovering relationships between multiple layers of information (Wang, 2006). This link begins with a process called geocoding, discussed further in Chapter Four. In short, geocoding computes spatial coordinates for a given point based on an assigned projection (data for this dissertation utilizes longitude and latitude). Once projected, a visual representation of the data can be displayed for analysis. With geocoding, GIS surpasses its traditional role of mapping to enable greater analysis through the statistical analysis of spatially associated data.

Many of the modern theoretical advances in GIS are attributed to cartographer and geographer Waldo Tobler (1970) and his law of geography which states "everything is related to everything else, but near things are more related than distant things" (p. 236). Because proximity constitutes a variable of influence, researchers may analyze the geographical distribution of a dependent variable to forecast trends and identify relationships (Tobler, 1970). Geographically speaking, the various combinations of social behaviors and spatial attributes are limitless, yet by mapping commonalities, visible patterns of attributes emerge (Robinson, 1959). The advantage is in the ability to provide a sense of place to "identify where similarities, correlations, and interactions exist" (Steinberg & Steinberg, 2006). In doing so, spatial disaggregation is possible, and

analysis becomes more comprehensive providing a richer understanding of the issues (Fotheringham, Brundson, & Charlton, 2002).

Complementing traditional regression analysis with spatial analysis enables researchers to visualize distributions of the students to identify enrollment patterns (Mitchell, 1999). Spatial analysis can take multiple forms, but at its foundation it is the search for patterns based on geographic inputs (Rogerson, 2009). With these geographically based patterns, one can assess the nature of patterns, or intensity of patterns (Mitchell, 2005). "Geographic patterns range from completely clustered at one extreme to completely dispersed at the other" (Mitchell, 2005, p. 75). If the observed pattern is between the two extremes, it would be termed a random pattern. With a null hypothesis related to the random distribution of cases in a given spatial area the pattern can be statistically analyzed to "calculate the probability that a pattern isn't simply due to chance" (Mitchell, 2005, p. 75).

Spatial Analysis in Enrollment Research

To date, a limited amount of literature has been published on spatial analysis applications within enrollment management, although there are strong indications this will change (Tate, 2008). Some of the earliest work relating to spatial analysis in the context of enrollment came from research determining optimal locations for new schools (Stewart, 1941). Stewart (1941) used geographic-based analysis of populations to establish market-based guidelines for establishing colleges. He used gravity-potential theory which is derived from Newton's gravitational energy law which identifies the

interaction effects that location and masses have upon gravitational force. According to Newton's Law, one mass exerts a force on every other mass relative to the distance and mass (Putnam, p. 124). Stated in planetary terms, "Jupiter is attracted by the Sun, as is Saturn, but Jupiter and Saturn also attract each other" (Laskar, 1995, p. 54). Stated in enrollment management terms, the institution (one mass) causes a desire to enroll (exerts a force of attraction) for a student (mass) relative to proximity, and the proximity of other students who are attracted to the institution (McConnell, 1965). Gravitational potential examines the influence of a population located at one point on a population located at another and can be stated mathematically as: $I_{1j} = P_1P_j/D^{x}_{i-j}$ or the "interaction (I_{ij}) between two populations (P_i, P_j) is directly proportional to the product of two populations (P_iP_j) and inversely proportional to the distance intervening between the points locating the two populations (D_{i-j}), or to that distance taken to some exponent (D^{x}_{i-j})" (McConnell, 1965, p. 29). Later this theory was applied to the spatial variability of college enrollments (McConnell, 1965; Kariel, 1968). McConnell (1965) found "the spatial distribution of enrollees by area of origin for an educational institution is essentially a spatial distribution of migrants attracted to an opportunity located at a point in space" (p. 30). He found that an institution's ability to draw students, decreases with distance, and that uneven distribution of enrollments could be attributed to per capita income levels, population density, and the student's proximity to other large universities (in Ohio) (McConnell, 1965). In a study on enrollments in Washington, Kariel (1968) found that proximity influenced behavior most when the costs (in geographic terms) were low, i.e. short distances from the institution, short travel time, and fewer associated travel

expenses. Later presentations and articles examining enrollment patterns supported the earlier findings of McConnell and Kariel (Marble, Mora, & Herries, 1995; Marble, Mora, & Granados, 1997; Mora, 2003; Smith, Spinelli, & Zhou, 2002). Other studies have used spatial analysis to identify recruitment areas providing more or fewer students than would be statistically expected (Zhou & Wu, 2005; Martin, 2001; Martin, 2002).

Recently, spatial analysis has been explored as a tool to study access, diversity, and policy implications for universities. Goldrick-Rab and Benson (2006) examined spatial processes in college access. By using spatial analysis, the authors were able to visually display predictors of college access to test whether these patterns were random or clustered (Goldrick-Rab & Benson, 2006). Additionally, one recent study examined how policy changes may impact admission applicants with respect to diversity, academic quality, and neighborhood income levels (Ousley, 2007). By mapping the location of applicants, changes in applicant characteristics could be correlated with institutional policy changes. Finally, the 2008 presidential address to the American Educational Research Association was titled *Geography of Opportunity: Poverty, Place and Educational Outcomes* (Tate, 2008). The expanded version of this address, published as an article, called for greater diligence on the part of researchers to consider the geographical context of data and the influence of geography on opportunity (Tate, 2008). Specifically, Tate (2008) wrote "it is critical that factors associated with human development and opportunity be salient features of the evaluative process" (p. 397). With past studies identifying the importance of proximity in relation to enrollment, the future application of spatial analysis in the field of enrollment management appears to

hold great relevance for assigning geographic context to existing literature, and for improving continuing research.

Enrollment Management

Spatial analysis provides greater understanding of the context of a given subject. In this dissertation, it provides a context for enrollment management, and so it is important to understand how enrollment management came into existence, and how it is utilized today. Institutions of higher education continue to evolve and reorganize in an effort to create management systems that maximize efficiency and effectiveness (Birnbaum, 2000). Enrollment management is one such adaptation. There is a belief among many in higher education that strategic enrollment management has become a necessity rather than a convenience (Hossler, 1984). Due to the complexity of overlapping and sometimes competing institutional goals, it is important to understand the historical development of the enrollment management office within higher education. The idea of enrollment management as a distinct entity emerged in the 1980s as two separate institutional units, admissions marketing and student retention, began to be viewed as interdependent in their contributions to the success of an institution (Kemerer, Baldridge, & Green, 1982; Hossler & Bean, 1990). One reason was that low student retention was seen to have its roots within the initial recruitment of students and the positioning of the institutional brand with respect to the student expectations marketing created (Miller, Bender, Schuh, & Associates, 2005). The strategic management of institutional enrollments became critical due to mounting external pressures to shape the composition of a class to achieve institutional goals. Public institutions of higher

education are expected to compose a class that maintains or increases high academic standards, retention rates, diversity of the student body, and affordability—all without increased reliance on state financial assistance (Quirk, 2005; Slaughter & Rhoades, 2004; Dolence, 1993).

Hossler and Bean (1990) described enrollment management as the institution's efforts to influence the demographic composition and size of the student body. "Organized by strategic planning and supported by institutional research, enrollment management activities concern student college choice, transition to college, student attrition and retention, and student outcomes" (Hossler & Bean, 1990, pp. 5-6). Others expanded its role to include "an institutional research and planning function that examines, and seeks to manage the flow of students to, through, and from college" (Kroc & Hansen, 2003). How inclusive enrollment management became on each campus, depended upon the individual institution's level of adoption. As opportunities arose to consolidate related responsibilities, enrollment management's responsibilities expanded, and for many campuses it now encompasses the entire student experience including student recruitment, campus visits, admissions, financial aid, orientation, advising, and registration (Hossler, 1984; Black, 2001). Enrollment management now has the responsibility for lifecycle management for students including "entrance, student persistence, and student outcomes" (Hossler, 1990, p.16). While the responsibilities have expanded, the ultimate goal has remained the same— to recruit graduates (Hossler, 1984; Hossler & Anderson, 2004).

Recruiting graduates at a public university is not as simple as having the desire, because the goal of improving student retention may inhibit the achievement of other goals. As an example, while the institution might improve retention rates through higher admission standards, it is a solution that could reduce the overall number of enrollees, net tuition revenues, and campus diversity (Quirk, 2005). Rather than experience the reductions, the enrollment manager may need to offset losses by increasing enrollments from nonresident populations (Hossler & Anderson, 2004). At public institutions, this reduction of in-state enrollments might adversely impact diversity goals, and might run afoul of state requirements regarding class compositions (Ehrenberg, 2000; ABOR, 2005). Public institutions are controlled by the states in which they reside and therefore have legal obligations to serve the citizenry. Various state mandated formulas exist to ensure the citizens' needs are met, while balancing the need for revenue generated from higher nonresident tuition (ABOR, 2005). Institutions strive to have a diverse student body, yet must balance this with the complexity of race considerations in admission decisions, consistent with federal and state laws (Gratz v. Bollinger, 2003). With each incoming class, enrollment managers must balance the intersection of social, political, and fiscal restrictions of the state. In *Tuition Rising*, Ehrenberg (2000) noted an enrollment management goal is to obtain the best students the university can attract, but enroll them "in a way that does not require it to devote an excessive share of the University's resources to financial aid" (p. 171).

The beginning of successful admissions and retention lies in the successful recruitment of desirable applicants. The definition of "desirable" is dependent upon the

enrollment manager's goals for a given class. The recruitment occurs pre-enrollment, while the retention efforts of campuses are most visible post-enrollment; however, enrollment management views the goals of these two activities as a single goal of graduating students. The reason is that persistence indicators are evident when initial contact is made with prospective students via electronic and direct mail solicitations. As a corollary outcome, the success or failure of a group of recruited students only can begin to be measured after one, four, or even six years have passed, providing quantifiable evidence as to those students who ultimately persisted at the institution of original enrollment.

Enrollment management: recruitment and admissions.

With an understanding of the goals of enrollment management, the discussion shifts to understanding admissions' recruitment as a method for achieving the stated goals of an institution. The idea of recruiting as a necessary function has been around since the 18th century (Hossler, 1984). Although modifications have occurred since its introduction, the basic understanding of recruiting and admissions as a series of events begins with Philip Kotler's description of the admission's process as a funnel. The admissions funnel involves potential students involved in five distinct stages going from the largest number of individuals (prospective) to the smallest (matriculants) (Ihlanfeldt, 1980; Litten, Sullivan, & Brodigan, 1983; Kotler, 1976). The five stages listed from largest to the smallest number of individuals are: Prospectives, Inquirers, Applicants, Admitted Applicants, and Matriculants (Kotler, 1976). During the various stages of the funnel, the demographic makeup of the class dramatically can be changed based on

decisions of the enrollment manager. Also, at various stages great disparities, and available choices, may exist between high-selectivity and low-selectivity institutions. Highly selective institutions traditionally have a large number of inquirers and applicants, yet have a proportionately lower number of admitted applicants (U.S. News and World Report, 2009). By default, this enables greater options related to crafting the class including selecting those with the most desirable characteristics including SES, academics, race and ethnicity, and geography (Hossler, 1999). Less selective schools may have a smaller number of applicants and applicants with less desirable characteristics. These schools may need to fill an equivalent number of seats (or greater due to higher attrition rates), and would be forced by market conditions to admit lower quality applicants to reach enrollment goals (Hossler, 1999; U.S. News and World Report, 2009). To mediate this effect, institutions attempt to have the largest pool of qualified applicants possible (Kotler, 1976). To improve the odds for having a large pool of potential students, the prospective student stage is critical to improving institutional characteristics related to the student body. For many years now, geodemography has been used at this stage to identify geographic areas containing students most likely to enroll at the institution (Hossler, 1999; Mora, 2003; Marble, Mora, & Herries, 1995; Marble, Mora, & Granados, 1997). Several higher education consulting companies offer fee-based multivariate analysis utilizing large geodemographic datasets to provide institutions with lists of students most likely to enroll. The industry leaders, Noel-Levitz' ForecastPlus, College Board's Recruitment PLUS, and the American College Testing Service's Educational Opportunity Service products focus on enrolling students with

various goals including increasing diversity and targeting special populations associated with gender, geography, family income, religion, majors of interest and race and ethnicity (ACT, 2009; College Board, 2009a; Noel-Levitz, 2009). Similarly, literature discusses the potential applicant pool along the same lines: test scores, grades, high school of attendance, church affiliation, miles from campus, or other specific geographical locations (Ihlanfeldt, 1980). However, neither the literature nor the consulting firms mention utilizing these data to identify students most likely to graduate (ACT, 2009; College Board, 2009a; Noel-Levitz, 2009; Ihlanfeldt, 1980).

A prospective student pool can reach into the hundreds of thousands dependent upon enrollment management goals and resources. As mentioned before, this pool often consists of the institution's primary target market. If successful outreach is made to an individual in the prospect pool, the individual responds to the institution's marketing efforts and is re-categorized as having moved to the inquiry stage. As the individual passes onto the next level of the funnel, additional information is gathered enabling greater accuracy in predicting probabilities. A student moving from prospect to inquiry, may supply the institution with an updated mailing address, e-mail address, major of interest, or other information increasing the enrollment manager's knowledge of the individual.

At the inquiry stage, the goal is to convince qualified students to apply for admission to the institution. Once the potential student has moved to the application stage enrollment managers again have the opportunity to shape the class. Additionally,

the application for admission and the accompanying submission of test scores and high school transcripts provides additional data to base an admissions decision upon. Mitchell Stevens writes admission officers now work in an "information –based evaluative regime" (Stevens, 2007). For admission decisions, applicants may be placed into one of three applicant pools: accepted, denied, and potential admits (Litten, Sullivan, & Brodigan, 1983). The first category, admits, includes those applicants who are academically well-qualified based on institutional standards. The second category, denials, includes those students who are not considered academically qualified for the institution. The third category, potential admits, offers enrollment managers the most opportunity, but provides the greatest risk for shaping the class because these students are somewhere in the middle with respect to their qualifications for acceptance (Camara & Kimmel, 2005; Stevens, 2007). Decisions regarding these potential admits are based on the enrollment goals and needs of the institution. If the institution seeks to increase enrollment, greater numbers from the potential admits pool are accepted. If the institution seeks to increase average test scores, or grade point averages, fewer applicants from this pool may be accepted. If the institution has accurate prediction capabilities related to retention, only those students most likely to graduate would be accepted if the goal is to improve retention.

Between the admitted stage and the matriculation stage (enrollment), the university still may lose admitted students. For this reason, enrollment managers must use research analysis when deciding on the number of students accepted. Of primary importance for this is knowledge of the institution's yield rate, that is the number of

enrolled students divided by the number of accepted students (Sullivan & Litten, 1976). The advantage is that based on previous year's data, the enrollment manager may estimate the number of admitted students who will enroll, enabling greater accuracy for predicting the size of the next year's class (Sullivan & Litten, 1976). When the admitted students arrive on campus at the start of the next semester, the job of admissions concludes, but the enrollment management responsibilities continue (Ingersoll, 1988).

Enrollment management: retention.

The previous discussion on recruiting and admissions addressed the difficulty of meeting institutional goals when recruiting an incoming class. Once an enrollment manager achieves these goals, retaining this mix of students until graduation becomes the primary responsibility for enrollment management through its student retention efforts (Hossler, 1984). While a certain level of attrition is expected (estimated by past enrollment/persistence trends), and is factored into each year's student recruitment, one year of recruitment success quickly can be undone through high attrition rates. During the 1990s, findings were released from three major longitudinal studies on persistence illustrating the severity of the problem at the nation's four-year institutions. These studies included the National Center for Education Statistics study (NCES), the Cooperative Institutional Research Project (CIRP), and the National Education Longitudinal Study of 1988 (NELS). Based on these studies, after four years of enrollment, 31-36% of enrollees earned a bachelor's degree from the school of initial enrollment (Adelman, 2006). After six years, 53-58% had earned a bachelor's degree from the school of initial enrollment (Adelman, 2006). This means that nationally, at

four year institutions, only 42-47% of the freshman class enrolling in 1989 earned a degree from the institution of initial attendance (Berkner, Cuccaro-Alamin, & McCormick, 1996; Adelman, 2006). In a separate study focusing on public institutions at various levels of selectivity (as defined by SAT/ACT scores of the entering class), 26-65% of enrollees were able to complete after four years, and 66-86% completed after six years (Bowen, Chingos, & McPherson, 2009). The explanation for the range disparity can be explained when the results are disaggregated by graduation rates at the less selective state institutions (26% graduated in four years, 51% in six years) versus the graduation rates at highly selective state flagship institutions (65% graduated in four years, 86% in six years) (Bowen, Chingos, & McPherson, 2009). At any level of selectivity, sustaining such dramatic losses in the student population leaves the institution with few options other than to recruit and enroll a greater number of new students at substantial costs (Litten, Sullivan, & Brodigan, 1983). This creates a vicious cycle where the enrollment manager attempts to maintain target enrollment levels, knowing that a predictable number of the new recruits will never graduate. By reducing the number of withdrawals the cycle can be broken; however, institutions have been attempting to break this cycle for nearly 100 years (Hossler, 1984).

This is not to say improvements in understanding retention issues have not been made. In fact, many early retention studies remain relevant today. One theory that evolved into other theories was William Spady's (1971) balance theory. Balance theory focused on the characteristics of the individual student and made the analogy that when a student does not complete their passage through the higher education system, that student

has committed a form of academic suicide (based on Emile Durkheim's sociological studies on suicide) (Spady, 1971). Vincent Tinto (1993) expanded Spady's work in combining elements of Durkheim's suicide theory with Van Gennep's theory on rites of passage into community membership. The result was one of the most influential (and debated) theories of student retention, known as the student integration model (Tinto, 1993). Utilizing this framework, the process of integration into a college community can be viewed as an interactive process where individuals act to reshape their environment.

Those critiquing student-integration theory (Tinto himself has critiqued this theory) have greatly expanded retention theory by noting three major shortcomings with the theory. This includes its failure to consider economic factors including financial aid, tuition and other economic factors (St. John, Cabrera, Nora, & Asker, 2000; Stampen & Cabrera, 1986; Stampen & Cabrera, 1988; Murdock, 1986; Nora, Attinasi, & Matonak, 1990). "Significant economic shifts, changes in student loan programs, unexpected changes in family and/or individual finances, and termination of part-time employment may all act to significantly reduce the available resources students have at their disposal for college attendance" (Tinto, 1993). When financial needs become a psychological stressor, they can divert the student's attention from academic to monetary concerns increasing the odds for withdrawal (Braxton, 2000). Now, most retention research includes at least some consideration of the economic factors influencing student persistence.

A second critique of Tinto's theory calls for a model focusing less on the social factors and more on the behavioral (psychological) factors influencing retention (Pascarella & Terenzini, 1991; Braxton, 2000; Bean & Metzner, 1985; Cabrera, 1990; Astin, 1993; Milem & Berger, 1997). One reason is that social and psychological factors become intertwined making it difficult to measure one set of variables without the other. "Poor academic performance can be a sign of low academic integration; and social, academic, persona, and emotional adjustment, all of which can be seen as related to social and academic integration" (Smedly, Myers, & Harrell, 1993; Baird, 2000, p. 71). The social and psychological aspects have greatly expanded understanding of the myriad of reasons students cease enrollment.

Finally, a third major critique referenced in this writing relates to a failure to recognize the importance of different cultures. Tinto's student integration theory does not adequately account for cultural differences between majority and minority student populations. William Tierney (1992) argues that because Tinto makes this error, and misuses the term "rite of passage," the theory of student integration does not provide a workable model of integration for the ethnic minority student. Instead of thinking about student participation from a social-integration perspective, Tinto's theory does not use "ritual" and "rites of passage" in the same context as Van Gennep who never used these terms to refer to movements between distinct cultures (Tierney, 1992). Van Gennep "never assumed that a Sioux youth underwent an initiation ritual in Navajo society. Yet Tinto's model assumes that same Sioux youth will undergo a rite of passage in Anglo society" (Tierney, 1992). Tierney advocates that we should "conceive of universities as

multicultural entities where difference is highlighted and celebrated" (Tierney, 1992). According to existing non-culturally-sensitive theories, a student would improve the chances for graduation through greater integration into the new institutional environment and fewer social interactions with the previous environment. In a sense, Tinto's integration theory advocated that the student must commit cultural suicide to embrace a new life in the institutional culture. If they fail to break from the old, they will not succeed in the new. "Though one may develop a strong attachment to the immediate group, one's attachment to the institution is likely to be considerably weaker" (Tinto, 1993). Tierney argues that it should not be assumed that ethnic minority students must abandon their culture to succeed in college. In later editions of his work, Tinto does clarify that majorities are made up of several communities that can include ethnic minority students. He states that membership only requires some membership within at least one community instead of total emersion into the new and total denial of the old community. Still, his theory of integration takes the student from their former culture into the institutional culture.

As Tinto viewed the problem, it was the individual's responsibility to integrate, going through the rites of passage until they fit into the new culture. The focus was on the individual who commits suicide/drops out, instead of being on the culture. Tierney argues that this desire to fix a "perceived problem" is functionalist thinking, and fails to address the world in which the individual exists. "Without struggling to understand how that world has been constructed we are doing little more than reinscribe notions of power

and privilege for those who have had an active hand in determining those very relations of power" (Kuh & Love, 2000, p. 216).

Beyond the three main categories of retention studies (economics, social and psychological, and cultural), it is important to consider the point in a student's enrollment that withdrawal occurs. Several researchers have identified the first year as being the most critical for new students (Choy, 2002; Ishitani, 2008). While first year retention is important, when the goal is to understand graduation, enrollment managers must remember that factors influencing year one withdrawal may be different than those effecting withdrawals in later years (Ishitani & DesJardins, 2002; Bowen, Chingos, & McPherson, 2009). A large proportion of students who leave between the first and second year transfer to another institution (Bowen, Chingos, & McPherson, 2009). When excluding these transfer students, the number of students withdrawing (and not enrolling at another four-year school) is more equally distributed since "nearly half (44%) of all withdrawals occur after the second year" (Bowen, Chingos, & McPherson, 2009, p. 35). When disaggregating by institution type, the less selective institutions experienced a higher proportion of year-one withdrawals than did the more selective schools.

Institutional Fit and Market Segmentation

To pursue this line of investigation, this research focuses on the pre-enrollment demographic and geodemographic characteristics of students to see how various characteristics influence persistence and graduation outcomes. The exact nature of these characteristics influencing outcomes is dependent upon the identity of the institution.

This intersection of the individual with the institution has become known as institutional fit, “the degree of congruency, or fit, between student needs, abilities, interests, and goals and the ability of the institution to adequately respond to those needs, abilities, interests and goals could lead to increased student satisfaction, academic achievement, and personal growth” (Williams, 1984). Institutional fit holds that the better the fit, the more likely a student is to persist (Williams, 1984). This idea has been around higher education for many years with institutions attempting to maximize the fit between students and the institution (Creager, 1968; Painter & Painter, 1982; Pace, 1980). Rather than trying to forcefully integrate students into a campus culture forcing them to lose self identity, institutional fit posits that a student’s fit within the institution can be estimated before matriculation. Because the basic tenets of institutional fit (also known as student-institution fit) are nearly identical to those of market segmentation, tenets of this related idea are presented. In marketing literature, individuals who have purchased a given product in the past have more in common with individuals who will purchase the product in the future, than with individuals who will not purchase the product. This marketing strategy seeks to “identify and delineate market segments” by dividing “total demand into relatively homogeneous segments which are identified by some common characteristics” (Tynan & Drayton, 1987, p. 301). “These characteristics are relevant in explaining and predicting responses of consumers, in a given segment, to marketing stimuli.” At its core, market segmentation seeks to identify “the most appropriate variable or variables with which to subdivide total demand into economically viable segments” (Baker, 1984, p. 123).

Institutions ignoring “fit” or a “misfit” and attempting to be all things to all prospective students, participate in mass marketing of the institution, i.e. “mass produced, mass distributed, and mass promoted… to all consumers in an attempt to obtain economies of scale” (Tynan & Drayton, 1987, p. 303). Targeted marketing through a market segmentation strategy seeks the opposite and attempts to identify students possessing characteristics of past graduates. “Instead of scattering their marketing effort, they can focus it on the buyers who have the greatest purchase interest” (Kotler, 1984, p. 251). Based on market segmentation and institutional fit, the pre-enrollment characteristics relevant to this study will be discussed with respect to the existing literature.

Institutional characteristics.

The characteristics of the institution account for many factors influencing persistence. This study focuses on a single institution, and while the institutional characteristics as such might seem homogenous, the ways in which the institution affects students would be different for each student within the analysis. This is why, as mentioned as a limitation, the exclusion of post-enrollment factors influences the outcomes of this research, and should be considered in interpretation of the results. Since there is evidence that institutional characteristics influence students differently, the literature regarding this side of the student-institution-fit equation with respect to persistence is mentioned. The first characteristic is the type of institution. Typically, students attending private universities persist to the second year, and graduate at higher levels than those attending public institutions without considering other potential

influences (Pascarella & Terenzini, 2005; Horn, 1998; Astin, Tsui, & Avalos, 1996; Berkner, Cuccaro-Alamin, & McCormick, 1996). However, when controlling for pre-enrollment characteristics of the students, there is little difference between persistence rates at the two institution types (Saupe, Smith, & Xin, 1999; Astin, Tsui, & Avalos, 1996; Lee, Mackie-Lewis, & Marks, 2003). Past studies also have considered the size and selectivity of the institution as they relate to persistence. It appears size of the institution has little to no influence on persistence although there is a small negative influence for Caucasian and Hispanic students (Astin, 1993; Astin, Tsui, & Avalos, 1996). Admissions' selectivity results in a higher persistence rate (Saupe, Smith, & Xin, 1999). However this may simply speak to selective schools' ability to "enroll students with higher academic abilities, better elementary and secondary school preparation, higher degree aspirations, clear and higher occupational goals, and greater family economic and other resources" (Pascarella & Terenzini, 1991). While post-enrollment institutional retention efforts undoubtedly influence persistence and graduation outcomes, they are not the focus of this study, and consequently are not included in the analysis.

Pre-Enrollment Characteristics Related to Persistence

The focus of this study is on the pre-enrollment characteristics of students, including characteristics that may identify students most likely to remain enrolled at the institution of study. For Hossler (2007), this selection of characteristics is about finding the appropriate institutional fit. Traditionally, this means admitting "only those students who have earned good grades in high school, who have high standardized test scores, and who are highly motivated" (Hossler, 2007, p. 104). In addition, Hossler (2007) noted

these students are most often from affluent and well-educated families. Research on the related issues of access, first year persistence, and graduation has identified the variables with the strongest influence as gender, race and ethnicity, high school coursework, SES, educational aspirations and family support, socioeconomic status, financial aid, precollege encouragement programs show the most significance (Berkner, He, & Cataldi, 2002; Berger & Milem, 1999; Elkins, Braxton, & James, 2000; Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006).

One of the most comprehensive studies of characteristics predicting student performance in college is from Cliff Adelman (1999; 2006) of the U.S. Department of Education. His, *Toolbox Revisited* examined data from the National Center for Education Statistics, NELS: 88/2000 study examining students from 8th grade through years that would include college graduation. This study is comprehensive in its scope and contains several variables not available to the enrollment management officer making decisions about admissions. Still, its breadth of analysis makes it an invaluable resource as a starting point for this study. Additionally, the work of Astin (1982), and Pascarella and Terenzini (2005) provided the foundation for variables included in this research.

Gender.

Gender appears to have little, to no influence on persistence when controlling for other contributing characteristics. One study utilizing NELS data, one of the most comprehensive data sets available to researchers, found that being male decreases the probability for degree completion by 11% (Adelman, 2006, p. 24). However once high

school academic history is included “demography plays a considerably reduced role” (Adelman, 2006, p. 24). Others could not find significance for the impact of gender with respect to degree completion (Alexander, Riordan, Fennessey, & Pallas, 1982).

There is strong evidence to suggest that gender will play a more important role in predicting persistence in the future. Already, the numbers of women enrolling has increased (U.S. Department of Education, 2008). Several studies have found women graduate at a slightly higher rate than men (Astin, 1993; Kanarek, 1989; Bowen & Bok, 1998; U.S. Department of Education, 2008). This may increase given the fact that high school girls regularly perform at higher levels than boys in “college preparation as measured by achievement test scores and by math- and science- course taking” (Goldin, Katz, & Kuziemko, 2006, p. 134). In addition to performance, the incentive to perform has also increased for females due to improved earning potential, greater labor market opportunities, and an increase in the age of first marriage (Goldin, Katz, & Kuziemko, 2006). Finally, females may be gaining a greater sense of entitlement to a college education, and may experience less pressure to fulfill traditional homemaker roles due to a change in societal attitudes toward the role of women in the workplace (Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006, p. 18; Astin, Oseguera, Sax, & Korn, 2002).

Race/Ethnicity.

By extension, numerous studies have found race/ethnicity to be an indicator of persistence due to the college readiness of students disaggregated by race (Fuller, Manski, & Wise, 1982). Factors impacting degree attainment begin before attendance, as

only 35% of African American and Hispanics enroll in college (Carter & Wilson, 1997). Once enrolled, ethnic minorities are less likely to graduate (Light & Strayer, 2002; Venti & Wise, 1983; Porter, 1990; Tinto, 1993). Being an ethnic minority reduces the probability of earning a bachelor's degree by 17% (Adelman, 2006, p. 24). Alexander, et al (1982) found a 10% reduction for African Americans versus Caucasian students. On a positive note for minorities, when controlling for other factors such as SES and gender, some studies have found ethnic minority students persist at higher rates than Caucasian students (Light & Strayer, 2002; Alexander, Riordan, Fennessey, & Pallas, 1982).

The institutional setting of race also may be critical factor. African American students on predominantly Caucasian campuses persist at lower rates (Astin, 1982; Thomas, 1981). When African American students attended a predominantly African American institution, or women attended a women's college, both groups were more likely to graduate versus attendance at a predominantly Caucasian institution, or coeducational institution (Pascarella & Terenzini, 2005).

One concern of past research was that using race as an admissions consideration has unfairly attributed poorer performance to racial attributes. This is due to the fact that ethnic minority students often experienced less advantaged upbringings than Caucasian enrollees (Vars & Bowen, 1998). Thus the persistence and graduation outcomes associated with race and ethnicity may have less to do with genetics, and more to do with the demography and the neighborhood disadvantages of low SES, low parental education, and other negative factors associated with persistence.

Recognizing the importance of race in the discussion of admissions, a corollary race-related variable derived from the census also will be used in the analysis. Using the percentage of a neighborhood composition that is ethnic minority in the analysis may reveal differences not readily noticed by analysis of race, regardless of other factors.

Residency classification.

A student's classification based upon residency has a twofold implication with respect to persistence. The first is that the student resides in a physical location that may be closer to the campus. Proximity can impact the availability of, and accessibility to, resources, i.e. ability to live at home, or to more easily interact with family and pre-college friends. Having an existing social structure in place may be an advantage to some students, but a disadvantage to other students trying to break away from their previous social group to integrate into a new set of peers (Tinto, 1993). The second implication of residency is the financial difference in tuition and fees paid. Resident tuition and fees during the four years of study ranged from $4,098 to $5,048, while nonresident tuition ranged from $13,078 to $16,282 (Fact book, 2007; Fact book, 2004). On average, nonresident students in this sample are more affluent. Should the student consider transferring to another institution, the nonresident would have more options due to this affluence. A resident may have college choices limited to two comparatively priced in-state universities, while the nonresident has many national options in a similar, or lesser, price range.

Urban versus rural.

Caison and Baker (2007) point out that past research using rural/urban classifications often use different definitions of these two categories, and caution should be used when comparing studies. Illustrating the point, the Census Bureau changed its definitions between the 1990 census and the 2000 census, recognizing a group of small towns previously classified as rural, share more commonalities with urban clusters (Caison & Baker, 2007). The changes resulted in a 3 million person decline in rural population compared to the 1990 census. Had the definition remained the same, the result would have been an increase of 2 million people for the rural population category.

There is a definite attainment gap when comparing urban and rural graduates. One study found rural students were less likely to attend college than their urban counterparts (Gibbs, 1998). Gibbs (1998) argues that it is a lack of access to college, rather than persistence that creates a rural/urban gap in educational attainment. Still other studies found varying results when the results of outcomes for rural students were disaggregated by gender. Rural females are more likely to be encouraged to pursue the traditional role for women which is "to maintain the home and raise children" (O'Quinn, 1999; Caison & Baker, 2007). Still, studies regarding the outcomes of students from rural or urban areas indicate there may be a small negative effect for rural students when it comes to persistence (Feller, 1974). It is believed the gap relates to resources. Echoing these findings, Blackwell and McLaughlin (1999) found that rural students had similar educational attainment as urban counterparts when they controlled for family, school, and community factors.

High school GPA.

When controlling for demographic variables, academic variables were "much more strongly related to college graduation than the student's background characteristics" (Alexander, Riordan, Fennessey, & Pallas, 1982). Academic variables for this study include high school GPA, and the highest standardized test score on either the SAT or ACT.

GPA in high school has been shown to be a positive predictor of first year retention (Mathiasen, 1984; Mouw & Khanna, 1993; Pike & Saupe, 2002). It is also highly predictive of scores on standardized tests (Noble, Davenport, Schiel, & Pommerich, 1999). As a predictor of first year retention, it is one of the pre-college characteristics that consistently predicts student performance in college (Pike & Saupe, 2002). Extending beyond the first year of attendance, GPA has been shown to be predictive of the student's odds to complete college (Astin, Tsui, & Avalos, 1996). Astin, Tsui and Avalos (1996) found students entering college with an A-level GPA were "four to five times more likely to finish college" than students with C-level or less GPA (p. 11).

SAT/ACT scores.

The institution of study no longer requires a standardized admissions test as part of the application process. Still, taking the test may be evidence of cultural capital as 97% of all enrollees took either the SAT or ACT prior to enrolling. Akin to high school GPA, standardized test scores are viewed as one of the most predictive academic

variables for predicting college grades, which directly correlates with persistence (Mathiasen, 1984; Pike & Saupe, 2002; Vars & Bowen, 1998; Hezlett, et al., 2001; Bridgeman, Pollack, & Burton, 2008). Others found a direct positive relationship between SAT scores and graduation (Bowen & Bok, 1998; Vars & Bowen, 1998). For Bowen and Bok, it is noted that "overall graduation rates achieved by these schools cannot be attributed solely to their success in enrolling students with high SAT scores" (Bowen & Bok, 1998, p. 63). Instead, they ascertained it was the school of attendance that was a better predictor, than SAT scores. The College Board released two similar reports on SAT-related research nearly twenty years apart, and both demonstrated the SATs value in predicting college graduation (Burton & Ramist, 2001; Wilson, 1983). Burton and Ramist (2001) found that the majority of studies found the combination of SAT scores and high school GPA proved to be the best predictor of college graduation. Astin, Tsui and Avalos (1996) found those in the highest SAT score interval were three times more likely to earn a bachelor's degree than those with a score below 700.

Standardized test scores and their predictive power must be considered in the context of race and ethnicity. A study of enrollees at highly selective institutions found significant differences in the graduation rates of African-Americans versus Caucasian students (Vars & Bowen, 1998). These differences were found at all levels of SAT test scores, and remained for both male and females (Vars & Bowen, 1998). When disaggregated by gender, SAT scores are better predictors for females than males, and females graduate at higher rates (Vars & Bowen, 1998).

High school coursework.

A variable commonly used with the research literature that is not used here is the coursework taken in high school. In 1992, those completing various levels of math courses in high school were scored on their probability of attaining a bachelor's degree. The results for coursework (percentages) were as follows: Calculus (83.3), Precalculus (74.6), Trigonometry (60), Algebra 2 (39.3), Geometry (16.7), Algebra 1 (7), and Pre-algebra (3.9) (Adelman, 2006, pp. 30-31). This variable may be more predictive than even test scores for predicting academic success and persistence (Gladieux & Swail, 1998; Adelman, 1999). While access to the course titles taken by individual enrollees was available from applicant transcripts, the creation of a compendium to evaluate all courses and codify coursework for the purposes of analysis was not feasible given the scope of this study.

First generation college student/family educational levels.

As higher education becomes more universal in the United States, the number of first generation students is increasing (Trow, 2000). Defined as students who have not had either parent attend college, first generation students are attending college at higher rates, but persist at lower rates (Choy, 2001). For 1992 high school graduates, parental education strongly influenced college enrollment—59% enrolled if their parents had no college, 75% enrolled when parents had some college and 93% when parents held a bachelor's degree or higher (Choy, 2001). However, only 21% enrolled in four year institutions versus 65% of high school graduates whose parents held a bachelor's degree

or higher (Choy, 2001). In 1995, first generation students accounted for 34% of all college students attending four-year institutions (Kojaku & Nunez, 1998).

Pascarella and Terenzini (1991) noted that social reproduction is one of the most beneficial impacts of college education—"intergenerational transfer of benefits from parents to children". Not having parents who have already experienced, and graduated from, college has a negative impact on one's chances for graduating (Choy, 2002, p. 22; Ishanti, 2006). Adelman (2006) found first generation students were 21% less likely to graduate than those whose parents had graduated from college (p. 23). Choy (2001) observed this phenomenon should not be a surprise as first-generation students "tended to report lower educational expectations, be less prepared academically, and receive less support from their families in planning and preparing for college than their peers whose parents attended college" (p. 10). Part of the reason for lower performance because "first-generation college students tend to be at a distinct disadvantage with respect to basic knowledge about postsecondary education (e.g., costs and application process), level of family income and support, educational degree expectations and plans, and academic preparation in high school" (Pascarella, Pierson, Wolniak, & Terenzini, 2004).

First generation students are more likely to withdraw after their first year, and less likely to obtain a degree. This in part attributed to the difficulties encountered by all college students, compounded by "substantial cultural as well as social and academic transitions" not experienced by students whose parents graduated from college (Pascarella, Pierson, Wolniak, & Terenzini, 2004, p. 250). Several other factors were

found to be more prevalent in first generation students including: likely to work more hours during college, less likely to live on campus, less likely to integrate into the campus culture, and they earn lower cumulative grades (Pascarella, Pierson, Wolniak, & Terenzini, 2004; Pike & Kuh, 2005).

The experiences of first-generation students are not always negative. Due to the deficit of capital, they tend to make stronger incremental gains than peers who are not first generation. In the vein of some being better than none, Ishanti (2006) found first generation students were more likely to persist when they had a higher GPA and parents who had completed at least some college.

Socioeconomic status/family income.

One of the most predictive variables of persistence and graduation is socioeconomic status. Past studies have utilized various combinations to define this metric. Among those included are: annual family income, father's education, mother's education, father's occupation, household items, Pell eligibility (Alexander, Riordan, Fennessey, & Pallas, 1982). The National Educational Longitudinal Survey (NELS) and High school and Beyond (HS&B) defined the socioeconomic status variable based upon a combination of family income, parental education levels and occupation (Carnevale & Rose, 2004). Bowen and Bok (1998) define SES by a combination of family income and parental education.

Studies have found a significant impact of high SES upon the probability of attaining a bachelor's degree (Alexander, Riordan, Fennessey, & Pallas, 1982; Titus,

2006). This is because the first challenge is for low SES students to enroll, as they are less likely to consider college as an option (Walpole, 2007). In fact, "Being from a lower-SES household is a far greater liability than being black" (Alexander, Riordan, Fennessey, & Pallas, 1982, p. 328). Additionally, other variables such as gender and race are more pronounced when combined with socioeconomic status (Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006, p. 18; Hamrick & Stage, 2004; King, 2000). Bowen and Bok (1998) found high SES students were predicted to graduate at a rate 8% higher than those from low SES. The difference for African American students was more noticeable, rising to a 15% difference between low and high SES enrollees with respect to predicting graduation (Bowen & Bok, 1998).

Related studies have found the negative effects of living in an impoverished area are magnified by corollary outcomes including lower quality schools, unavailability of suitable marriage partners, and reduced exposure to mainstream social networks and conventional role models (Wilson, 1992). A 2007 study found concentrated neighborhood disadvantages had a negative influence on college aspirations among African American adolescents (Stewart, Stewart, & Simons, 2007). When considering the evidence that lower SES results in a lower odds of graduation, it is important to note African American and Hispanic students in the High School and Beyond study were nearly three times as likely to come from families with incomes less than $20,000 (Kane, 1998).

Summary

Using spatial analysis informed by enrollment management admissions and retention literature, provides a foundation for greater understanding of the issues. Plotting a potential student's location in geographic coordinates enables the enrollment manager to merge these data with additional layers of detail, as in the case of U.S. Census data. The use of these two sources opens up new opportunities for social science researchers to explore the influence of neighborhood variables.

While using pre-enrollment characteristics from spatial analysis is the focus of this study, the use of student characteristics known to predict persistence must be used as controls. GPA, standardized test scores, rural/urban, gender, and race must all be incorporated to ensure a correct interpretation of results. Understanding the intersections of a student's characteristics and associated neighborhood characteristics enables this study to benefit both theory and practice. Finally, the presentation of these student characteristics provides a context for the discussion of theoretical framework, specifically how these characteristics fit with the theories guiding this study.

CHAPTER THREE: THEORETICAL FRAMEWORK

In Chapter One, the potential for using GIS within enrollment management to expand our understanding of the student characteristics influencing persistence and graduation was discussed. In Chapter Two, I presented literature to establish the foundation for this research. Emphasis was placed on the definition and use of spatial analysis, and how it may be beneficial for the enrollment management applications of admissions recruitment and retention. Additionally, literature on the use of GIS for enrollment and educational research was discussed. In this study, I seek to further understanding as to the ways in which a student's spatial origin may inform persistence studies. To that end, this chapter introduces the theoretical constructs informing this study. I begin the discussion with a definition of capital. Immediately after I introduce the concept of human capital, and provide a discussion of its strengths and weaknesses. Next, I discuss the forms of capital, namely economic, social, and cultural capital, centering on the work of two of the most prominent researchers: Pierre Bourdieu and James Coleman. By focusing on these two theorists, the forms of capital will be discussed in terms of applicability to higher education research and this specific study. By using the forms of capital as a theoretical framework, I use this research to demonstrate how these various forms of capital serve as proxies for the selection of the student characteristics used in this study.

Finally, as a unifying framework, the theory of social reproduction is introduced as the mechanism for explaining how these forms of capital transfer from a neighborhood (including the parents, peers, and other social networks) to the enrolling students. By

using these conceptual frameworks, I will articulate the transference of persistence-related capital from a spatially identifiable locale (as measured by associated neighborhood demographics) to the enrolling student.

Theory of Capital Defined

The starting point for this chapter is with the theory of capital as derived from Marx (1849, 1933), who defined capital as a surplus value. In its basic form, capital refers to "assets that yield income and other useful outputs over long periods of time" (Becker, 2008, p. 1). In this "classical theory of capital," capital is both the "product of a process" (resulting in a surplus value) and "an investment process in which the surplus value is produced and captured" (Lin, 1999, p. 29). This aspect of capital is economic capital, that is the wealth that a class possesses that is "immediately and directly convertible into money and may be institutionalized in the form of property rights" (Bourdieu, 1986, p. 243). The investment of the surplus value is made by the dominant class with the expectation of a return in the form of increasing surplus value (Lin, 1999). Since Marx introduced his ideas on capital, others such as Weber (1968), Salisbury (1969), and Durkheim (1973), have built upon and expanded the concept of capital to include various nuances. While it is beyond the scope of this writing to elucidate all scholarly contributions to the understanding and development of capital, several major concepts, including human, economic, social, and cultural capital, are presented below.

The Theory of Capital Expanded: Human Capital Theory

Human capital theory states that expenditures on items such as education or training are investments in capital, but are unique because they produce human assets, instead of financial or physical assets (Becker, 1964). In Marx's concept, capital served as a way to exploit the social relations between the classes for the benefit of the dominant class; however, human capital took capital from simply an economic investment in inanimate objects to an investment in that which creates a surplus of knowledge within humans (Schultz, 1961; Becker, 1964; Becker, 2008). With human capital, no longer was the value realized only by the capitalist class, but by the individual laborer who accumulates a surplus value as well in the form of knowledge, skills, health, and values (Schultz, 1961; Becker, 1964; Becker, 2008; Lin, 1999). Through this theory, the investment in individuals is viewed as a form of capital—"the labourer is himself a produced means of production, an item of capital equipment" (Johnson, 1960, p. 562). Human capital theory would predict that a student choosing to invest in a college education would reap a well paying job, or contentment through a fulfilling career (Becker, 1964).

However, human capital theory is problematic since it would predict that this investment in education by an individual would result in reciprocal benefits derived from the investment regardless of class, and that "academic success or failure is an effect of natural aptitudes" (Bourdieu, 1986, p. 47). Indeed, its weakness rests in its economic origins, as the investment rests on monetary investments and profits—quantifiable factors

that are convertible into money. Bourdieu (1986) notes human capital theorists fail to explain,

> different proportions of their resources which different agents or different social classes allocate to economic investment and cultural investment because they fail to take systematic account of the structure of the differential chances of profit which the various markets offer these agents or classes as a function of the volume and the composition of their assets (p. 48).

Human capital lends itself to the study of student characteristics used in this study including GPA, gender, race, and best standardized test scores. Still, it does not lend itself to explaining differences between students with similar backgrounds and would not predict imputation of effects from neighborhood characteristics, other than economically quantifiable characteristics such as household income levels. Given the literature reviewed in Chapter Two, it seems misguided to rely solely upon human capital theory when previous research has concluded the existence of class differences and their influence on persistence within higher education (Pascarella & Terenzini, 1991; Choy, 2001; Ishanti, 2006). This study will analyze neighborhood demographic characteristics that are the result of class differences, and an appropriate framework must accommodate their inclusion. Given the premise of this dissertation is to explore how these characteristics impact persistence and graduation probabilities, human capital is not a suitable framework without additional theories of capital.

Bourdieu's Forms of Capital

While human capital does not meet all the needs of this study, it was introduced because the economic aspects of the theory fit well with portions of the research question. Also, the idea that individuals could possess this surplus addressed an important need required by the selected framework. For these reasons, the discussion moves to a theory of capital that embraces aspects of human capital, while further expanding the idea of capital. It was Bourdieu who believed that an economic-only view was severely limiting for understanding the ways in which humans function within the social world. As a response, Bourdieu (1986) put forth that economic capital needed to be viewed as a form of capital that existed along with two other forms of capital—cultural capital and social capital. To more fully understand this view, a discussion of cultural capital and social capital, along with their value to this study, is provided.

The Forms of Capital: Cultural Capital

Bourdieu (1986) defines cultural capital in terms of its value, it is "convertible, on certain conditions, into economic capital and may be institutionalized in the form of educational qualifications" (p. 47). Similar to human capital, there is an investment by the dominant class for the purpose of reproducing a set of symbols and meanings such as education (Bourdieu & Passeron, 1977; Lin, 1999). Here, the cultural capital can include "general cultural awareness, knowledge skills, and verbal facility that are acquired in childhood socialization" (Swartz, 2002, p. 552). Once obtained, this form of capital may be passed from one generation to the next (Bourdieu & Passeron, 1977). The method of this transference from parent to child or teacher to student is critical to understanding its

innate value, and is made possible by various forms cultural capital can take. Bourdieu (1986) viewed cultural capital as existing in three forms: embodied, objectified, and institutionalized state. The first, embodied cultural capital, is a form of capital that is physical in nature, and obtainable only by the individual's exertion of effort toward the goal of self-improvement (Bourdieu, 1986). While this form of capital cannot be purchased, there is an associated cost—time. In many societies, time is considered a luxury item, and for this reason, embodied cultural capital is associated with wealth. It is not immediately transferred, but instead accrues over time through repeated exposure. It is basically socialization to behave in a socially desired manner. Examples of this type of capital could include the ability to engage in educated conversation, having knowledge of socially acceptable behaviors and topics, knowing what to write in an admissions application essay, or knowing how to converse in an admissions interview. The amount of time required for its accumulation takes longer than most forms of cultural capital as it is gained during the entirety of one's socialization. For this reason, possession of embodied cultural capital merits value through its scarcity, and increases one's prestige within a society (Bourdieu, 1986).

The second form of Bourdieu's cultural capital is the objectified state and is found in "cultural goods," the evidence of this would be found in objects within the home such as a piano or other instrument, artwork, a computer with educational software, or books in a home library (Bourdieu, 1986, p. 47). In a sense, these objects provide the means to obtain cultural capital, while at the same time providing tangible evidence of one's potential for acquiring cultural capital. As the parents provide children with the

opportunity to obtain cultural capital, the transference from one generation to another generation becomes possible.

The third and final form of Bourdieu's cultural capital is the institutionalized state. The institutionalized state is a form of objectification conferred on an individual by an institution in the form of an educational degree, or licensure thus guaranteeing the transference of cultural capital (Bourdieu, 1986). This capital can be seen in the ways in which middle-class and upper-class families value education as a means of ensuring the economic success of their children (McDonough, 1997). In this case, the transference often is possible through the affluence of the parents who pay for the education, or the society which pays through subsidizing public education, or the student who sacrifices other needs and wants to obtain the desired capital.

Bourdieu's Cultural Capital in Higher Education Research

With an understanding of Bourdieu's definition of cultural capital, it is important for this research to discuss how cultural capital informs higher education research. The advantages of Bourdieu's (1986) view of cultural capital are that it enables the analyses of stratification and inequalities. Because cultural capital provides access to scarce rewards such as a college education, it allows for the measurement of capital as a quantifiable resource that is transferrable between generations (Lareau & Weininger, 2003). Cultural capital informs our understanding of the relationship between one's social class and one's educational attainment (Bourdieu & Passeron, 1977; Bourdieu, 1986; Giroux, 1981; Valadez, 1993; London, 1992). Without this, or a similar

framework, one might look at educational attainment as an outcome of meritocracy. One might expect the smartest students, those who study the hardest, and those with the most dedication to persist and graduate, while those without these characteristics would fail to persist (if they even make it to the enrollment stage) (Trow, 1984). However, Hearn (1991) examined the college destinations of high school graduates and found that while academic factors influenced college choice the most, nonacademic factors (i.e. cultural capital) played a role and should not be overlooked. McDonough (1997) found student's cultural capital determined not just college destinations, but also where students thought they could attend, in effect creating class-based college aspirations. DiMaggio (1982) found significant positive effects of cultural capital on attendance and on the predicted outcomes of enrollees (DiMagio & Mohr, 1985). These are just a few of the applications of cultural capital within higher education literature, and the numerous other studies cannot be presented in this limited space. However, as these and other studies have evidenced, cultural capital and its recognition of class-based factors enable higher education research to explore the relationship between class and attainment, making it well-suited for this research.

The Forms of Capital: Social Capital

To understand what social capital is, this discussion begins with the work of Bourdieu (1986), and then moves to the definition offered by James Coleman (1988). In a broad sense, social capital is a form of capital about relationships. Although often combined, cultural capital and social capital are distinct entities. The confusion results from the fact that social capital is difficult to assess through quantitative variables since it

is defined as an individual's connections with others (Bourdieu, 1986). Instead, it is often the evidence of social capital that may be analyzed. Bourdieu (1986) defines social capital as an "aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition—or in other words, to membership in a group" (p. 248). He adds that this membership provides the members with "the backing of the collectivity-owned capital, a 'credential' which entitles them to credit in the various senses of the word" (Bourdieu, 1986, p. 248).

Coleman's concept of social capital.

While Bourdieu's definition of social capital begins to enable this study to articulate differences based on neighborhood characteristics, he views social capital as a negative force allowing the dominant class to suppress the lower classes (Dika & Singh, 2002). Coleman's view instead takes a positive view of social capital as a form of social control where "trust, information channels, and norms are characteristics of the community" (Dika & Singh, 2002, p. 34). This view helps build a framework of understanding to explain impacts of neighborhood characteristics on student persistence and graduation. Coleman (1988) states social capital should be defined by its function, not by any one descriptor, but by a variety of inputs. Instead of being a student's action that one could measure, it is the resource for that action (Coleman, 1988). The student obtains this resource through the normal social relations, group membership and family interaction providing them with informational advantages (Portes, 1998). This is critical to this dissertation, because a suitable theoretical framework must enable the observation

of the sources of social capital, or at least allow the quantification of its existence. But, as mentioned, social capital is not an action, only a resource for said action. If social capital is "embodied in the relationships among persons" as Coleman (1988) states, how can it be measured? The solution is to examine those resources through a lens that allows for the analysis of the action (outcomes). Since social capital is seen as being derived from relationships within one's social networks, choices leading up to an action such as enrollment would be expected to be influenced by the prospect's peers including parents and those within the social structures (Portes, 1998). Social capital is a form of capital found within the relationships. By socializing, beneficial behaviors are transferred and learned. Because of this social transference, this study can examine the characteristics of those comprising the students' social circle—families, neighbors, mentors, and peers (Pascarella, Pierson, Wolniak, & Terenzini, 2004, p. 252; Coleman, 1988). Differences in social structures, be they families, neighbors, mentor or peers, should result in observable differences in persistence outcomes. As a result, students with greater social capital resources to draw from would be expected to obtain greater amounts of social capital and persist at higher rates (Coleman, 1988).

Coleman's social capital in higher education research.

Coleman's theory of social capital enjoys widespread use as an explanatory theory for educational attainment in K-12 and higher education. Dika and Singh (2002) assert that the vast majority of K-12 education literature may cite Bourdieu (1986), but upon inspection they instead utilize Coleman's framework. Musoba and Baez (2009)

suggest the same is true for much of higher education research in the past, underscoring the importance of Coleman's theory within higher education research.

Past research has found that social capital acts as a tool by which the students may best take advantage of available resources (Coleman, 1988; Portes, 1998; Morrow, 1999). Prospective students may utilize this capital to maximize their opportunities and make effective decisions, and it remains a resource for improving the prospective student's chance to improve personal and economic status (Coleman, 1988; DiMagio & Mohr, 1985; Lamont & Lareau, 1988). Coleman used a social capital framework as a way of looking at how a child's family background impacts educational outcomes (1988). Since Coleman, others (in addition to the literature in Chapter Two) have used capital to examine college enrollment. Perna & Titus (2005) found a student's odds of enrolling in higher education increased with parental involvement. Hossler, Braxton, & Coppersmith (1989) and Hossler, Schmit, & Vesper (1999) found that relationships and social networks (parents) have a critical role in the decision making process of students regarding college attendance. Similarly, Zweigenhaft (1993), McDonough (1997), Freeman (1997), and Hurtado, Inkelas, Brigs, & Rhee (1997) noted access and opportunity for different races and ethnicities was strongly influenced by cultural capital factors. And once in college, the cultural capital continues to influence students as Carbonaro (1998), Paulsen & St. John (2002), and Longden (2004) all found cultural capital was an important factor in determining whether or not a student persisted.

The Forms of Capital: An Integrated Framework

Using capital as the theoretical framework and specifically economic, cultural and social capital enables this dissertation to connect student characteristics with the research. A person possessing economic wealth may not possess the cultural capital to know how to best utilize their resources. A first generation student may not know the importance of obtaining an education—information a second generation student would be more likely to possess. These informational advantages may be in the form of taking the ACT or SAT multiple times to achieve a better score, or they may have an association with submitting an application for admissions early, and to several schools. For these reasons, student characteristic variables evidencing capital are included in this study.

The Reproduction of Capital

The final piece of the framework deals with the means for the transfer of capital from one group to the student. For this to occur, individuals must be exposed to culture over time through social exchange (Bourdieu, 1986). The process of socialization is known as social reproduction (defined as the transfer of capital from one party to another), and ensures that social class structure persists from one generation to the next (Bourdieu & Passeron, 1977). According to Bourdieu (1984, 1986), capitalist societies reproduce the dominant class through the transference of capital from one generation to the next (Bourdieu & Passeron, 1977). In this view, educational attainment is one device used to pass along existing class qualities—and inequalities. Evidence of this can be seen in a correlation between socioeconomic status and degree completion (Stevens, 2007; Bowles & Gintis, 2002; Bowles & Gintis, 1976). For Coleman (1988), it is less a method

for domination, and he includes the idea with social capital—making transference part of the social capital accumulation. To pass along this knowledge, one must already be in possession of the knowledge— the Marxist notion that the educator must himself be educated to pass along knowledge. The transference of the educator's knowledge to the student is the activity that creates a "profound and lasting transformation…to the extent that they are prolonged in an action of continuous inculcation" (Bourdieu & Passeron, 1977, p. 32). This passage is not guaranteed to occur for all individual's but is influenced by class, racial position, and an individual's characteristics (Lareau & Horvat, 1999).

The concept of social reproduction theory accounts for much of the educational and status attainment gaps between social classes within higher education (McDonough, 1994; DiMagio & Mohr, 1985). One way this occurs is through schools rewarding students with the highest levels of cultural capital (Bourdieu, 1977; DiMaggio, 1982). Evidence of cultural capital can be found in the parents of students who are more likely to possess higher levels of education, a group that includes teachers, doctors, lawyers, and executives (Bourdieu, 1984; Bourdieu & Passeron, 1977; Madigan, 2002). As a result, this creates a vicious cycle whereby those with the greatest resources reap the greatest rewards such as better grades, promotion, and ultimately graduation (Hearn, 1991; DiMaggio, 1982). Enrolling students possessing the greatest cultural capital would lead to greater success with respect to improving persistence (Berger & Milem, 1999; Milem & Berger, 1997; Berger, 1997). It also would lead to reduced access for those not possessing the appropriate amount of cultural capital. Those with the lowest probability of being exposed to cultural capital and subsequently acquiring it are those with the

lowest amounts of education (Madigan, 2002). A life of pre-enrollment cultural deprivation, would translate into a lack of opportunity for low cultural capital students to succeed. Educational institutions perpetuate the reproduction and the inequalities by rewarding "the inherited cultural capital of the dominant classes by implicitly requiring it for good academic performance and systematical devaluing [sic] the culture of the subordinate classes" (Swartz, 2002, p. 553). The disadvantages already experienced by those without capital are reproduced by the structure of college admissions are reproduced (Hearn, 1990). Viewing retention from a social reproduction perspective requires an examination of both the individual factors and the organizational factors (Berger, 2000). Based upon the reproduction, the expectation is that variables evidencing cultural and social capital will translate into a higher probability for academic success (Milem & Berger, 1997). Therefore, many of the variables included in this study will be selected based on the premise that cultural capital and social capital are reproduced (Bourdieu, 1986; Coleman, 1988). With these reproduced attributes, variables indicating the holder originates from an environment possessing high educational credentials imply membership within the dominant class. And with these attributes, "dominant class groups are able to capitalize on that advantage by translating their inherited cultural resources into high levels of scholastic performance and attainment" (Swartz, 2002, p. 553).

Conclusion

This chapter presented theories of capital to provide a framework for the data analysis and findings of this dissertation study. By beginning with a definition of capital,

the emphasis was placed on making an investment to achieve a surplus value of outcomes. Followed by the idea of human capital, a framework embracing that this capital could exist in both the dominant class and the laborer began to emerge. By looking at Bourdieu's cultural capital, the framework accommodated the idea that capital was not only about economic capital or monetary resources, but could take the form of reproducible symbols and meanings. Finally, with a discussion of Coleman's theory of social capital, the framework was in place to understand how social capital was not a negative form of class domination, but served as a positive form of social control, where order was achieved through replication of desirable individual characteristics. In this way, this study will view student characteristics as proxies for the various forms of capital. By placing these student characteristics within the framework of economic, cultural and social capital, this study will analyze the transfer of capital from parents, peers, and the neighborhood to the enrollee—an example of social reproduction evidenced by student and neighborhood characteristics. It is my premise that cultural, social, and economic capital will help to explain student persistence and graduation for the sample used in this dissertation study.

CHAPTER FOUR: RESEARCH METHODS

Overview

This chapter reviews the research methods employed in this study. First, the procedure of secondary data analysis is presented with a discussion of its strengths and weaknesses. Next, the secondary data sources, including U.S. Census data and institutional data are discussed, along with missing data. The process of geocoding is presented as both a way to join the two data sets and to facilitate spatial analysis. From these data, the dependent variables and independent variables selected for this study are described. The data analytic methods employed, t-tests and regression analysis, are presented next to address the study's research questions:

- What are the geospatial patterns of enrollment for the institution of study?
- When using geospatial analysis, what are the characteristics associated with a student's probability to persist?
- When using geospatial analysis, what are the characteristics associated with a student's probability to graduate?

Finally, to conclude this chapter, a discussion of the limitations of this study is offered.

Secondary Data Analysis

This research will utilize data from multiple secondary data sources, defined as data not gathered to answer the specific research question, nor gathered by the researcher (Stewart & Kamins, 1993; Frankfort-Nachmias & Nachmias, 1992; Fortune & McBee, 1984). This is different from primary data, defined as data that involves the researcher being responsible for the "design of the research, the collection of the data, and the analysis and summary of the information" (Stewart & Kamins, 1993, p. 3). I made the choice to use secondary data based on several advantages it provided. The first advantage to using secondary data was the efficiency provided by its immediate availability (Cowton, 1998). Since the data already were collected, surveys did not have to be created, distributed, collected, or coded (although recoding was necessary) (Hakim, 1982). In this case, the data have already been manipulated, problems identified, solutions addressed, and weights provided (Carter, 2003). Having these tasks completed by others, increases the efficiency of research expenditures and allows for the targeting of research dollars to gaps in the knowledge (Stewart & Kamins, 1993). In addition to time, efficiencies are experienced with respect to the cost of obtaining data. Often, especially in the United States, governmental bodies produce large amounts of data and make it available to the public (Cowton, 1998). As a public record, the cost of these data provide a major advantage—free to the user, although vendors charge for cleaned data that provide usable file formats for the researcher (Stewart & Kamins, 1993). These advantages influenced my decision to use secondary data.

Data

To investigate the spatial distribution of enrollees and geographically related student characteristics associated with persistence/graduation, this study (as mentioned) will combine data from two secondary data sources. The first source is data from U.S. Census Bureau, and specifically from the 2000 U.S. Census (referred to hereafter as census data). The second source of data is the institution's office of enrollment management, which provided data collected for of all first-time, full-time undergraduate students enrolling for the summer and fall semesters of 2004 through 2007 (referred to hereafter as institutional data).

Census data.

For this research, secondary data from the 2000 U.S. Census were selected for analysis. This research uses the CensusCD 2000 data supplied by commercial vendor GeoLytics, and contains the entire US Census Bureau's SF3 dataset. This dataset contains 5,500 variables at the block group level for the United States (GeoLytics, 2010). The 2000 U.S. Census utilized two physical forms for data collection. The short form was provided to nearly five out of every six households. The short form "contains population questions related to household relationship, age, sex, relationship, race, Hispanic origin, and tenure (i.e., whether home is owned or rented) (U.S. Census Bureau, 2003a). Additionally, the long form was provided to approximately one out of every six households, and contains all of the questions from the short form, plus "additional detailed questions relating to the social, economic, and housing characteristics of each individual and household" (U.S. Census Bureau, 2003a). Since the 2000 census, the long

form is no longer in use, and now detailed socioeconomic data are collected annually through the American Community Survey (U.S. Census Bureau, 2007a).

The most detailed level of categorization occurs at the block level (U.S. Census Bureau, 2003a). As defined by the census bureau,

> Census blocks are areas bounded on all sides by visible features, such as streets, roads, streams, and railroad tracks, and by invisible boundaries, such as city, town, township, and county limits, property lines, and short, imaginary extensions of streets and roads. Generally, census blocks are small in area; for example, a block bounded by city streets. However, census blocks in sparsely settled areas may contain many square miles of territory (U.S. Census Bureau, 2003a).

For each census block, corresponding data are collected on the residents within the boundaries. However, due to their small size, generally about 85 people, long form data are not released for census blocks (Peters & MacDonald, 2004). As a direct result, a census block must be joined with other census blocks before it is publicly made available for analysis. This provides the second most detailed level of categorization, known as the block group (Figure 1). The size of a block group ranges from 600-3000 individuals, with 1,500 being the optimal level (1,000 is the optimal level on Native American reservations) (Stewart, Stewart, & Simons, 2007). The file used for this study contained 208,790 census block groups. Geographical boundaries of block groups are subject to change as population shifts warrant. Because these groups provide an aggregate compilation of data, long form information is available to the public.

Using the data from the 2000 U.S. Census enables access to vast amounts of data on residents of the United States. The scope of coverage makes this an invaluable

resource. Also, the depth of coverage is a major advantage. As mentioned, 5,500 variables (resident/neighborhood characteristics) are provided. Categorically, these characteristics include income, housing, employment, language spoken, ancestry, education, poverty, rent, mortgage, commute to work, etc. (GeoLytics, 2010).

Figure 1: Census block group boundaries near The University of Arizona.

Institutional data.

The data set incorporated summer session enrollees to include students reported to the Integrated Postsecondary Education Data System (IPEDS), as freshman domestic applicants for full-time enrollment. As the institution encourages students from underserved populations to attend summer academic programs to increase retention rates, applicants for summer sessions were included in the analysis. Excluded from this dataset are international enrollees; enrollees without a permanent U.S. residence (including enrollees with non-US military addresses); enrollees entering with 24 or more hours of college credit earned through dual-credit, or attendance at community colleges while in high school; and enrollees without a valid address. This dataset will differ from official university-reported data due to the exclusion, and inclusion, of selected populations as referenced. Comparisons to official university-reported data yield slightly different results with respect to mean grade point averages (GPA), mean standardized test scores, and retention. Larger differences are notable in the reported number of enrollees.

Missing data.

Since the focus of this study is the use of geodemographic data and U.S. Census data to predict outcomes, several additional classifications of enrollees were excluded from the data set because the associated cases would not yield sufficient or accurate data, thus reducing the final sample size and distribution. Prior to the data export from university records, international addresses were removed. This removed the majority of international students. However, several international students provided U.S. addresses as their permanent residence. The majority of these addresses were geographically close

to the institution. To avoid data corruption, the university's identification of the students' as either domestic or international was used to remove all international students from the analysis. As a group, these students had slightly higher persistence rates; therefore, the exclusion lowered the overall retention rate for the institution. For similar reasons, the analysis also excludes U.S. residents of Guam, Northern Mariana Islands, Puerto Rico and the U.S. Virgin Islands, along with U.S. enlisted personnel using an APO address outside the U.S. as their permanent.

The census information used in this study is based on the address self-reported to be the student's permanent home address. It was not possible to match 106 student addresses with U.S. Census data, resulting in the exclusion of these student records. This excluded group had a first-year persistence rate of 72%, seven percentage points lower than the institutional average.

The use of a post office box as a permanent address poses an accuracy issue for geocoding because the postal recipient does not necessarily live in the block group associated with the address. However, numerous rural localities rely upon post office boxes as a centralized distribution point for mail. This is most noticeable on rural Native American reservations. Of the 624 Native American enrollees at the institution of study, 18% used a post office box for their permanent address. Given the importance of research inclusive of this racial category, addresses using a post office box were retained. It is of note that the first-year retention rate for those with post office boxes regardless of enrollment year was 33%, substantially lower than the institutional rate.

As mentioned earlier, of the 23,743 enrollees with associated geographical information. From this number, 864 were excluded due to missing information regarding a specific race. Of these, 517 were excluded on the basis of their decision not to identify themselves based on race and— selecting "choose not to respond" on their application for admission. Another 68 were excluded because they indicated their race as "other", but failed to provide further details. The remaining 1,041 records were located in block groups that contained data, but I deemed the data insufficient for use in this study. These were student records in a block group with a median household income of zero dollars. Upon inspection, these addresses were within nonresidential areas, such as a university. Because no households exist within these block groups, they do no yield a median household income. Given the inaccuracy of the median household income level, it follows that other student characteristics associated with the block group may also yield misleading data. From the original sample size, 92% of the records had complete student characteristic data resulting in a final size of 21,838 records (Table 1). The group of 1905 records missing data had a one year persistence rate of 79%, similar to the institutional rate during this time period.

Table 1
Missing data

Year	Missing	(%)	Not missing	(%)	Total	(%)
2004	593	(11%)	5043	(90%)	5636	(100%)
2005	414	(7%)	5435	(93%)	5849	(100%)
2006	489	(8%)	5396	(92%)	5885	(100%)
2007	409	(6%)	5964	(94%)	6373	(100%)
Total	1905	(8%)	21838	(92%)	23743	(100%)

Note. The percentages may not add up to 100% due to rounding.

Sample

To answer the research questions mentioned at the beginning as they relate to a single institution, the sample was derived from The University of Arizona, a major public research extensive university in the southwestern United States. This institution enrolled 27,379 undergraduates from all states and 125 countries during the final year of this four-year study (Fact book, 2007). The University of Arizona admits approximately 80% of all applicants, allowing for it to be classified as either a selective institution or as a moderately selective institution depending upon the classification criteria (OIRPS, 2008; Bowen, Chingos, & McPherson, 2009; U.S. News and World Report, 2009). The sample used in my research contains all full-time freshmen enrolling (fall semester) at The University of Arizona from 2004 to 2007.

Measures

Dependent variables.

The first dependent variable used in this study, year-one retention (RY1), is based on the students' attendance in the year following their first year of enrollment. The second dependent variable, graduation (RY4G), represents graduation from the institution after four years of study, or continued attendance in the fifth year of enrollment. This information was obtained from the institution's office of institutional research. I obtained these data for all students enrolled as freshman from 2004-2007. A student attending in the fall semester of 2004 was later checked for attendance as of the 21st day of classes of the next enrollment year. In this case, if the student was enrolled in classes in the fall semester of 2005, the variable for retention in year one (RY1) was set to 1. If the student was not enrolled, the variable was set to 0. For the graduation variable, a student beginning in the fall semester of 2004 was later checked for graduation or enrollment as of the 21st day of classes of the 2009 enrollment year. In this case, if the student had graduated or was enrolled in classes in the fall semester of 2009, the variable for graduation in year four (RY4G) was set to 1. If the student was no longer enrolled and had not graduated, the variable was set to 0.

Independent variables.

Many of the independent variables are obtained from testing services or are provided by the potential student. For some students, the first official communication with institutions of interest begins in October of their junior year of high school. At this time the Preliminary Scholastic Assessment Test, commonly known as the PSAT, is

administered by the College Board. This test is proctored at public high schools and is optional for all students (College Board, 2009b). Prospects paid $12 for the test with fee waivers available based on financial need (College Board, 2009b). As one benefit, the test promotes itself as a valuable resource for those wanting "to receive information from colleges" by signing up with the Student Search Service (College Board, 2009b).

Following this test, institutions have the opportunity to purchase lists based on various criteria. A specific test score is not reported, but institutions may purchase lists on the "basis of a range of scores" (College Board, 2009b). Information sent to institutions includes: name; address; sex; birth date; school; grade level; ethnic identification (if provided); and intended college major (if provided). Institutions may request this information disaggregated by GPA, test score, intended major, geographic region or zip code (College Board, 2009b).

Other avenues for prospect-initiated contact include direct contact at recruitment events, high school counselor referrals, college fairs, and online inquiries including Web inquires and electronic mail. All of this information allows the institution to know that the prospective student has an interest, culminating in the submission of an application for admission. At this stage, the enrollment manager has a wide range of variables to assess the students' probability for graduation. Depending upon the time the applicant submits the application, variables assisting in the prediction of college retention may be known six to 12 months prior to the time of enrollment. Selected variables known at this time are discussed below and include the following: gender, race/ethnicity, residency

classification, standardized test scores, high school grade point average (GPA), and high school coursework (dependent upon the timing of transcript submission and review).

The independent variables for this study include student characteristics containing demographic information drawn from the application for admission, supporting documents sent to the institution at the applicant's request, and census data. As the choice of variables impacts and directs the entire research project, one goal of this chapter is to describe student characteristic variables selected for this dissertation. The challenge is to "strive for maximizing a model's ability to statistically account for variation in behaviors" and to attempt to "clearly explain the origins of particular types of disengagement behaviors" (Tinto, 1993). The goal was not 100% prediction of persistence outcomes since no theoretical model can explain "the wide range of behaviors that constitute the universe of social interactions" (Tinto, 1993). Instead, the goal is to inform the literature regarding spatial analysis and its application in persistence studies.

The student characteristics selected for this study are organized into three categories: students' personal characteristics, academic characteristics, and characteristics serving as proxies for capital (Tables 2-4).

The last category, proxies for capital, is disaggregated further to distinguish between economic, cultural, and social capital variables.

Table 2
Values for students' personal characteristics

Category	Variable name	Values
Gender	Gender	0 = male, 1 = female
Race/ethnicity	Caucasian	0 = no, 1 = yes
	Hispanic	0 = no, 1 = yes
	Asian American	0 = no, 1 = yes
	African American	0 = no, 1 = yes
	Native American	0 = no, 1 = yes
Location	In-state tuition (resident)	0 = no, 1 = yes
Location	Percent of block group classified as urban area	0-100
Location	Percent of block group classified as urban cluster	0-100
Location	Percent of neighborhood that is an ethnic minority	0-100
Location	Distance from university	0-3,049 miles

Note. Percent of neighborhood that is ethnic minority is also used as percent of neighborhood that is specific to the race/ethnicity being analyzed in the regression, i.e. percent of neighborhood that is African American, Asian American, Caucasian, Hispanic, or Native American.

Table 3
Values for students' academic outcomes

Category	Variable name	Values
Academic	HS GPA	0-4.0
Academic	Best standardized test score (ACT/SAT)	400-1,600

Table 4
Values for students' characteristics serving as proxies for the forms of capital

Category	Variable name	Values
Cultural	Application submission	0-14 (months before school)
Cultural	Number of standardized test submitted to school	0-12
Social	Percent of males in block group with 4-year-degrees	0-100 percent
Economic	Median household income level for the block group	$0-$200,000

Gender.

As mentioned, literature has identified the potential influence of gender warranting its inclusion in this analysis (Adelman, 2006; Alexander, Riordan, Fennessey, & Pallas, 1982). Self-identification of gender was requested on the institution's application for admission in each of the four years. Although optional, the variable existed on all records used for this study. In cases where the student may not have indicated originally, supporting documents including interest forms, or submitted ACT or SAT test scores were consulted to establish a gender. Available selections were "male" or "female" for this variable.

Race/Ethnicity.

A substantial volume of literature addressed race/ethnicity considerations in student persistence (Light & Strayer, 2002; Alexander, Riordan, Fennessey, & Pallas, 1982; Carter & Wilson, 1997; Pascarella & Terenzini, 2005). Given the multitude of race

and ethnicity combinations available through various sources including four years of admissions applications, standardized test scores, and self-reported sources, race and ethnicity were categorized by multiple terms. I found multiple terms were used within the data to describe race and ethnicity. To standardize the reporting, I used The University of Arizona's method for establishing six main race/ethnicity categories for domestic students. The categorizations and the included terms comprising the categories are the following:

- African American
    - African American
    - Black
- Asian American
    - Asian Indian
    - Chinese
    - Filipino
    - Hmong
    - Japanese
    - Korean
    - Laotian
    - Guamanian
    - Chamorro
    - Middle Eastern
    - Other Asian
    - Other Asian American
    - Other Pacific Islander
    - Samoan
    - Thai
    - Vietnamese
    - Native Hawaiian
    - Puerto Rican
- Caucasian
    - White
    - Caucasian
- Hispanic

  - Central American
  - South American
  - Cuban
  - Hispanic
  - Mexican or Mexican American
  - Mexican or Mexican American or Chicano
  - Other Hispanic or Latino
  - Other Hispanic, Latino, or Spanish Origin
- Native American
  - Alaska Native
  - American Indian
  - American Indian or Alaska Native
- Unknown
  - Choose not to respond
  - Other Race
  - Missing

Reflecting the multiracial background of many students, when multiple options declaring race/ethnicity are available, students at the university often chose more than one to accurately identify their race/ethnicity. In such cases, a single category was selected to simplify analysis. In the case of The University of Arizona, a hierarchy prioritizes domestic enrollees based upon ethnic minority representation with the most difficult populations to attract to the institution being first in the list. Race/ethnicity is assigned according to the following hierarchy.

1. When race is Native American enrollee is classified as Native American.

2. When race is Hispanic, and is not Native American, enrollee is classified as Hispanic.

3. When race is African American, and is not Native American, or Hispanic, enrollee is classified as African American.

4. When race is Asian American, and is not Native American, Hispanic, or African American, enrollee is classified as Asian American.

5. When race is Caucasian, and is not Native American, Hispanic, African American, or Asian American, enrollee is classified as Caucasian.

6. When race is unknown, and is not Native American, Hispanic, African American, Asian American, or Caucasian, enrollee is classified as unknown.

Once the race/ethnicity was determined per the hierarchy, variables were coded into six dummy variables (variables with a value of either zero or one) representing each race/ethnicity category.

Residency status.

The student characteristic of residency is included due to its potential influence on persistence as identified by literature on social integration and affluence (Tinto, 1993). Residency was based on the institution's categorization of the student for tuition purposes. The University of Arizona determines residency based on the state's board of regents policy. Residency requirements during the four years of study required students, or their parents/guardian/spouse to meet one of the following requirements:

- A resident of the state for at least one year;

- A member of the armed forces of the United States and claim the state as their legal residence for one year;
- A member of an Indian tribe recognized by the United States Department of the Interior whose reservation land lies in the state, and live on the reservation;
- A public school district employee under a full-time contract (ABOR, 2009).

High school grade point average.

With high school GPA identified by several past studies as the strongest indicator of college performance/persistence, it is included in this study (Alexander, Riordan, Fennessey, & Pallas, 1982; Mathiasen, 1984; Mouw & Khanna, 1993; Pike & Saupe, 2002; Astin, Tsui, & Avalos, 1996). High school GPA is initially provided by the student on the application for admission. This later is verified by the admissions office through the required submission of high school transcripts. Given that the application for admission may occur prior to the completion of the student's first and second semester of the senior year of high school, the GPA may not include grades from this time period.

Standardized tests (SAT/ACT scores).

The inclusion of standardized test scores is based on literature relating high test scores with persistence (Mathiasen, 1984; Pike & Saupe, 2002; Vars & Bowen, 1998; Hezlett, et al., 2001; Bridgeman, Pollack, & Burton, 2008). Scores for the SAT may range from a low of 200 to a high of 800 for each of the three test sections (College

Board, 2008). This study uses the composite score derived from the mathematics and critical reading sections of the test. The results of the writing section will not be included in this analysis. The range of standardized test scores analyzed was 400 to 1600. To achieve a single test score for analysis, multiple conversion methods were used depending upon the available scores. If a single SAT score existed, this score was used. In cases where an ACT test score was available in the student record, the ACT-SAT Concordance Tables provided by ACT were used to convert this score to an equivalent SAT score (ACT, 2007). These concordance tables were based on 2004, 2005 and 2006 data, making them ideal for converting test scores submitted to the university during the same time period (ACT, 2007). If this was the only score available, it was used for analysis. If multiple SAT tests existed in the student record, the highest available critical reading score was combined with the highest available mathematics score to provide a composite score. In all cases, the highest composite score, after comparing ACT and SAT composite scores, was utilized for analysis.

Additional student characteristics—proxies for capital.

While the preceding variables are known and exist as part of the institution's applicant record, several variables selected for this research have a predictive value, but are not available through transcripts, submitted test scores, or applications for admissions. The decision to include these additional variables was made based on existing literature of related variables and the theoretical framework. As the framework looks to see the impact of economic, cultural, and social capital as commuted through social reproduction, the following variables were selected as proxies for the various forms of

capital. Through geodemographic data collection, aggregate values of these variables based on the student's neighborhood of origin are available before the student submits an application. To explore the feasibility for obtaining, and applicability of using, these variables as derived from geodemographic data collection, this study incorporates these variables into the analysis.

Economic capital.

As mentioned, a challenge for enrollment managers is that few SES-related variables are available at the time of recruitment. Although educated guesses can be made based on high school of origin, the earliest an institution obtains student financial data is when (if) a Free Application for Federal Student Aid (FAFSA) is filed by the applicant. The FAFSA becomes available to the student in January of each year, eight to nine months before fall enrollment, but at a time when the majority of admission applications for the university have been submitted. While available in January, students may not fill out and submit the FAFSA until much later in the year.

Median household income.

However, through geocoding (discussed later in this chapter), aggregate variables associated with a student's economic capital become available for analysis. Although data are less student-specific than a FAFSA, these data provide an additional resource for predicting the persistence rates of prospective students. In this study, a neighborhood's median household income as reported in the census will be used as proxy for economic

capital. The median household income within a census block group was disaggregated into five groups based upon values. Values for income ranged from 0-$200,000.

Cultural and social capital.

First generation college student/family educational background.

The institution of study did not collect first generation information on the admissions application for all four enrollment terms analyzed, and as a result this specific variable is not included in this study. Given its importance in persistence literature, many institutions, including the institution of study now ask for the student's family educational background on the application for admission (Pascarella & Terenzini, 1998; Choy, 2001; Adelman, 2006; Ishanti, 2006). However, by including a variable derived from the U.S. Census neighborhood profile, I approximate a student's family educational level based upon educational attainment levels of those within the student's neighborhood of origin. In this study, data relating to the neighborhood educational attainment levels was disaggregated into 37 categories. The large number is due to disaggregation by the population as a whole, along with male and female populations.

The main disaggregations are:

1. No schooling completed
2. Nursery to 4th grade
3. 5th and 6th grade
4. 7th and 8th grade
5. 9th grade
6. 10th grade
7. 11th grade
8. 12th grade, no diploma
9. High school graduate (includes equivalency)
10. Some college, less than 1 year
11. Some college, 1 or more years, no degree
12. Associate degree
13. Bachelor's degree
14. Master's degree
15. Professional school degree
16. Doctorate degree

In accordance with the literature, three variables were constructed: percent of neighborhood population with a bachelor's degree or higher; percent of neighborhood population of females with a bachelor's degree or higher; and percent of neighborhood males with a bachelor's degree or higher. I then executed a correlation to determine the variable (of the three) most related to persistence (Table 5). In all years, the percent of neighborhood males with a bachelor's degree or higher showed the strongest correlation with first year persistence. This variable also showed the strongest correlation with students graduating or enrolled after Year Four.

Table 5
Correlations for block group residents with bachelor's degree or higher

Correlations to gender	Female	Male	All
Retained year one (RY1)	.055**	.064**	.061**
Graduated or enrolled after year four (RY4G)	.079**	.093**	.089**

** $p < .01$

Urban versus rural.

From the literature, originating from either a rural or urban background (and associated interactions with other student characteristics) has been shown to influence persistence (O'Quinn, 1999; Caison & Baker, 2007; Blackwell & McLaughlin, 1999; Gibbs, 1998). Through the use of census records, I will ascertain the degree to which a neighborhood can be classified as rural or urban. These designations are not dichotic, and a neighborhood (block group) often consists of more than one subcategory, and may consist of all four subcategories (U.S. Census Bureau, 2002). The neighborhood classifications of rural and urban are taken from the 2000 U.S. census.

The U.S. Census Bureau (2000) classifies as urban all territory, population, and housing units located within urbanized areas (UAs) and urban clusters (UCs). It delineates UA and UC boundaries to encompass densely settled territory, which generally consists of:

- A cluster of one or more block groups or census blocks each of which has a population density of at least 1,000 people per square mile at the time, and
- Surrounding block groups and census blocks each of which has a population density of at least 500 people per square mile at the time, and

• Less densely settled blocks that form enclaves or indentations, or are used to connect discontiguous areas with qualifying densities.

Within this urban category, two subcategories exist. The first is an urbanized area (UA) and "consists of densely settled territory that contains 50,000 or more people" (U.S. Census Bureau, 2000). The second subcategory, is an urbanized cluster (UC) and "consists of densely settled territory that has at least 2,500 people but fewer than 50,000 people" (U.S. Census Bureau, 2000). A rural designation provides the second major category and is defined as an area that does not meet the requirements to be defined by the urban classification standards. It is further disaggregated by farm and non-farm residents (U.S. Census Bureau, 2002).

For analysis, three variables were constructed: total urbanized population (comprised of UC and UA); total urbanized area population (UA); and total urbanized cluster population (UC). The total rural population was not used because its value is reciprocal to the total urban population, and only 9% of The University of Arizona sample lives in a block group with greater than a 50% rural population. Because the correlations for UC and UA demonstrated an influence in opposite directions, I decided to use both individually instead of using the composite urban variable (Table 6).

Table 6
Correlations of year-one retention for block group residents by urban and rural classifications

Urban/rural classification	Correlation
Urban (UA and UC)	.034**
Urbanized area (UA)	.040**
Urbanized cluster (UC)	-.022**
Rural (farm and nonfarm)	-.034**
Rural farm	-.035**
Rural nonfarm	-.010**

** $p < .01$

Geocoding the Data for Analysis

To create a common dataset for analysis and to obtain the student characteristic variables discussed above, I joined aspatial student data from the institution to the spatially dependent census data. This process required the conversion of each student's self-reported permanent address into spatial coordinates through a process known as geocoding, "geographical encoding the data to allow it to be mapped" (Jardine & Teodorescu, 2003). Geocoding computes spatial coordinates for a given point based on an assigned projection allowing a user to convert common street addresses into a meaningful coordinate system, such as latitude and longitude (Longley, Goodchild, Maguire, & Rhind, 2005). By associating these geographical coordinates with an individual's address, variables from the census can be appended to the student's record. To accomplish this, geocoding relies upon the joining of data from at least two databases. The first is the record containing an address and the second is the record "representing

the geometry of street segments between consecutive intersections, and the address-ranges on each side of each segment" (Longley, Goodchild, Maguire, & Rhind, 2005, p. 125). Based on a linear interpolation within the address range, a coordinate is estimated (Longley, Goodchild, Maguire, & Rhind, 2005).

For this study, a free online geocoding service, batchgeocode.com was used to provide longitude, latitude of the student address and the Euclidian distance from The University of Arizona's main campus in Tucson. This online service utilizes the *Yahoo!* Geocoding API which allows it to derive results from "a much more powerful commercial street dataset, one well known as an industry standard for transportation data" (Holmstrand, 2006). With coordinates for each address established, a layer may be created plotting each address onto a map. Without additional information, the addresses (plotted as dots) provide little detail beyond a general image resembling the United States (Figure 2). However, once the location has been established with coordinates such as longitude and latitude, these coordinates can be merged with the census data, providing details regarding the neighborhood characteristics associated with individual enrollees.

Figure 2: Spatial distribution of The University of Arizona enrollees 2004-2007, Continental United States. Substantial overlap of student addresses (stacking) occurs on this map due to scale.

Data Analysis

To answer the three main research questions mentioned earlier, this study utilizes spatial analysis and logistic regression analysis.

Spatial analysis

I will answer the first research question relating to the spatial distribution patterns of enrollments by using spatial analysis. This method will provide a visual representation of the student characteristics, illustrating geographically based enrollment patterns at the institution of study (Mitchell, 1999). Spatial analysis can serve as an exploratory technique for data analysis that relies upon the power of human intuition to make sense of visual patterns (Goodchild & Janelle, 2004). After mapping the general patterns of enrollment, I will examine the spatial patterns of features evidencing high probabilities for persistence and graduation. The *features* in this study will be the enrolled students. Once identified, underlying student neighborhood characteristics visually can be examined by displaying the patterns.

Rather than simply projecting dots onto a map, this analysis employs spatial *data* analysis. This method refers not to simple querying of spatial data, but also the use of statistical theory and techniques to model spatially-referenced data (Bailey & Gatrell, 1995; Krivoruchko, Gotway, & Zhigimont, 2003). Spatial analysis will be used to relate the location to data, enabling extraction of data related to the specific location of enrolling, persisting, and graduating students. These data can then be used as part of the second analysis method, logistic regression analysis. Following the use of these data in

the logistic regression, the resulting odds ratios may then be used to map the probability that a block group will contain students likely to persist at the university.

Logistic regression analysis.

With the census and institutional data merged into a common data set, the student characteristic variables referenced earlier in this chapter were selected for this study. In recent years, the statistical research method of choice for studying retention within higher education has been logistic regression. A survey of three major higher education journals from 1988-1999 found that 38% (greater than any other method) of all articles utilized a variation of regression (Peng, So, Stage, & St. John, 2002). The trend is increasing within higher education literature related to retention—attributed to the logistical regression's ability to explain relationships between a categorical outcome variable and a mixture of continuous and categorical predictors. This phenomenon does appear directly related to retention research, as a similar study focusing on educational research in general found only 12% of studies (on a variety of educational topics) used logistic regression (Hutchinson & Lovell, 2004). Logistic regression appears appropriate to specifically address research problems involving complicated data and categorical outcome measures such as enrollment/matriculation and retention (Peng, So, Stage, & St. John, 2002).

The strength of a logistic regression is related to situations where the researcher attempts to predict an outcome or identify the presence/absence of a characteristic or outcome (SPSS, 2002). Different from a linear regression, it is best is suited to questions

where the dependent variable is binary or dichotomous (Hosmer & Lemeshow, 2000). In the case of enrollment, the student either selects to continue enrollment or cease enrollment, thus lending itself to be identified by a dummy variable of one or zero. This study will utilize the dependent variables of RY1 and RY4G depending upon the availability for each cohort. These analyses will allow the regression coefficients to be used to estimate odds ratios for the independent variables (SPSS, 2002).

The logistic regression relies upon two basic assumptions. The first is the dichotomous relations, in that the distribution is associated with a binary outcome. The second is that there exists a relationship between the outcome variable and the independent variable(s), and this relationship can be accounted for by a logistic function (Cabrera, 1994). A regression can serve to predict or explain what has already occurred. Within higher education research literature the goal is most often to explain (Ethington, Thomas, & Pike, 2003). Theory shapes the selection of variables to be included within the regression. "Independent variables and the dependent variable used in the statistical analyses are directly specified by some underlying theory" (Ethington, Thomas, & Pike, 2003). Methods mean little unless integrated within a theoretical context, and in explanatory research data analysis they serve to shed light on theory (Pedhazur, 1982).

Based upon this foundation, this study will use a binary logistic regression to identify student characteristics influencing the student's probability to persist to the second year of enrollment. I will use gender, race/ethnicity, and residency status as my categorical variables. To find the variables most closely associated with the dependent

variables, the binary logistic regression will employ a forward stepwise method. This choice enables the consideration of a greater number of variables in the analysis. The reason is that the stepwise method in a logistic regression allows for control over how independent variables are entered into the regression model. Using a forward selection (Wald), variables will be entered into the equation one at a time. At each step, a single variable is added to the equation and tested for appropriateness to the model based on the significance of the score statistic, and the Wald statistic (a test of significance of the regression coefficient) (Wald, 1944; SPSS, 2002). A low Wald statistic probability results in the variable being excluded from the model, while a high Wald statistic probability results in the variable being retained in the model (SPSS, 2002). This choice will output two tables, one indicating variables significantly influencing persistence (retained in the model) and the second indicating variables found to have no significant impact upon persistence for the given model (excluded from the model).

Goodness of fit.

In evaluating the usefulness of a regression model, it is important to assess how well the model works, and how strongly it predicts the dependent variable (Menard, 1995). With the binary logistic regression using dependent variable values of either 1 or 0, it also is important to know "the frequency of correct as opposed to incorrect predictions of the exact value of the dependent variable" (Menard, 1995). To understand the model in terms of those correctly predicted, I use classification tables. These tables are available as an output of the SPSS software used in this analysis. These tables provide information on how well the model predicts outcomes. The output provides a

percentage of cases correctly predicted from the sample. This number can then be compared to the percentage of cases that would be predicted by chance. To determine the difference between the model and random chance, Menard (1995) recommends measuring the proportional change in errors. By measuring the proportional change in errors, fit of the model is established. By comparing the two predictions, I will determine whether the model improves prediction of the dependent variable. I will present this measure for all regressions in this study.

Limitations of this Study

This study of persistence and graduation has several inherent limitations related to secondary data, spatial data, logistic regression, and institutional data. This study was conducted at a single institution, and examined students enrolling during a four year period. As such, the findings are limited to interpretation based on the limited scope of the data set.

Missing data

As a result of the design of this research and missing data, the final sample size accounted for 92% of the student population enrolled during the years selected for this study. Because the purpose of this study was to analyze the impact of spatially related student characteristics, students without a valid address within the United States were dropped from this study. By default, this would disproportionately affect students using military addresses. Also affected would be students who did not enter their address correctly at the time of application. These categories were dropped from the data set used

in this study. Students not identifying race, or lacking any other student or neighborhood characteristic were dropped from the data set, resulting in a smaller final sample. For these reasons, the data may not be representative of the institution's original data.

Limitations of secondary data.

Because secondary data have already been collected, I needed to expend additional time becoming familiar with the nature of the data collection and how it was organized and formatted (Cowton, 1998). Without these considerations, "there is a danger that the researcher will misuse the data, perhaps drawing unwarranted conclusions" (Cowton, 1998, p. 428). The secondary data may be biased, and care is warranted in determining the nature of the data (Stewart & Kamins, 1993). Of course, this full impact of this problem may be difficult to assess if the collection methods have not been well documented (Frankfort-Nachmias & Nachmias, 1992). In my case, the census data have been well documented, and I possess first-hand knowledge as to the collection procedures involved in obtaining the institutional data from students.

Still, the secondary data have been collected for other purposes, and presented the disadvantage that it may not be specific to the research question at hand (Carter, 2003). This effects outcomes and theoretical considerations, as Cowton (1998) notes, "the researcher's attempts to manipulate the data into a suitable form, having been generated for another purpose they are likely to address less adequately than desired the theoretical concerns of the researcher" (p. 428). This is also the case for variables used in this study.

Since the datasets were created for other purposes, and not for my study, I was not able to choose the variables (or proxies) that fit perfectly with my theoretical framework.

A final disadvantage of secondary data is that it is rarely available at the individual level meaning the data have been aggregated. This was the case for census data, but not for the institutional data. The census data also presented the concern about the age of the data. By the time secondary data have been collected, analyzed for its original purpose, and released as a secondary source, it has aged and may no longer be accurate (Stewart & Kamins, 1993).

However, given the breadth of the U.S. Census, many of its disadvantages are related to its scope —the entire population of the United States. Due to this, considerations should be given to sampling, especially for ethnic minority populations. For the census data, an undercount occurs when people are not counted, and thus not included in the final counts. This may occur as a sheer undercount, defined as the "failure to count people who live in the nation" or it may occur as a differential undercount, "when some groups are undercounted more than others" (Peters & MacDonald, 2004, p. 17). Peters and MacDonald (2004) note certain populations have a greater risk of being undercounted versus other populations. For the 2000 census, the U.S. Census Bureau estimated an overcount of .49% for the entire U.S. population with a standard error of .20 and falling within a 95% confidence interval (2001; Peters & MacDonald, 2004). While percentage errors may seem small, the effects are magnified because the survey includes 280 million people. For selected populations the error rate

may be higher, disaggregated by geographical areas race and ethnicity, or living arrangements. These and other categories may have above average undercounts or overcounts (U.S. Census Bureau, 2003b). Table 7 provides an example of overcounts (shown as a positive number) and undercounts (shown as a negative number), as reported by the U.S. Census Bureau (2003b).

Table 7
Values for Census Bureau overcounts and undercounts by race/ethnicity.

Race/ethnicity	Under/overcount	Standard error
Caucasian	1.31%	(0.20)
Hispanic	-0.71%	(0.44)
Asian American[a]	0.75%	(0.68)
African American	-1.84%	(0.43)
Native American (On)[c]	0.88%	(1.53)
Native American (Off)[d]	-0.62%	(1.35)
Hawaiian[b]	-2.12%	(2.73)

Note. [a]Does not include Pacific Islander. [b]Includes other Pacific Islanders. [c]Includes only those living on a reservation. [d]Includes only those living off a reservation.

The use of census data also presents a limitation in that it assumes the U.S. Census Bureau accurately defined the boundaries representative of each neighborhood. The issue is that because census boundaries often follow topographic features (i.e. roads, rivers, mountains, etc.) these boundaries often divide areas with homogenous socioeconomic characteristics (Brack & Martin, 1989). Williams (1999) describes a census boundary as an area "approximated by neighborhood constructs," instead of "a place of cultural identify" (p. 1). While these units are established with consideration of

social, demographic, and historical attributes, the ability of such a unit to accurately represent the social and cultural identity of a community is limited (Williams, 1999; Zimmerman, 1994; Bowen, Salling, Haynes, & Cyran, 1995). Census designations of neighborhoods may be viewed as a crude and perhaps blunt unit of measurement for the true construct of the neighborhood. Still, given the national scope of this analysis, I consider census data to be the best available option for analysis of this sample.

Limitations associated with findings for ethnic minorities.

Although this research intentionally does not account for institutional characteristics in the analysis, interpretation of the results regarding ethnic minority populations should be considered within the context of campus demographics (Hurtado & Carter, 1997). Caucasian students account for 71% of The University of Arizona's student population in this sample. Meanwhile, Hispanic students account for 16%, Asian Americans 7%, African Americans 4%, and Native Americans 3% of students in this sample. Smedley, Myers, and Harrell (1993) concluded ethnic minority students on predominantly Caucasian campuses were more likely to experience stress related to their minority status. Because of this stress, ethnic minority students on a predominantly Caucasian campus such as The University of Arizona would be expected to have greater issues adjusting to life on campus, potentially resulting in lower year-one retention rates and lower graduation rates (Hurtado, Carter, & Spuler, 1996). Ethnic minorities on campus could experience greater difficulties relating to both academic and social areas of their college experience (Hurtado, Milem, Clayton-Pederson, & Allen, 1996). Members of ethnic minority populations encountering a hostile racial environment could be

expected to have difficulty adjusting to campus in nearly every facet, directly impacting their general sense of belonging and their chances to persist and graduate (Hurtado, Carter, & Spuler, 1996; Hurtado & Carter, 1997). Given the demographic composition of this sample, it is possible that the results would be biased for ethnic minority populations comprising a small percentage of the population on campus. Due to the design of this study, data are inconclusive regarding the exact nature of the effect that the sense of belonging for ethnic minority students attending The University of Arizona exerts on the findings. This is a known limitation of this study, and must be a consideration in any interpretation of my findings.

Limitations associated with spatial analysis.

Since the address matching process used for spatial analysis relies upon estimation, the level of accuracy is directly related to an assumption that addresses are uniformly spaced (i.e. 123 West Main Street is near 125 West Main Street), and to the accuracy of the database utilized (Longley, Goodchild, Maguire, & Rhind, 2005). Also, depending upon the method employed during data collection (secondary data outside my control), the completeness of the address, the newness of the address (i.e. a home in a 200-year-old neighborhood versus one in a new housing development) and the dataset, the accuracy of the output varies. As mentioned, with any secondary data, the age of the database used for geocoding must be considered (Stewart & Kamins, 1993). Geographical regions experiencing rapid growth result in data issues caused by new streets and new housing developments. Since this may distort the information, this has the potential to introduce anomalies of this nature.

Two types of data matching were available for my data set, with each influencing the number of matches and the degree of accuracy to a different degree. Deterministic matching uses "a combination of algorithms and business rules to determine when two or more records match (the rule 'determines' the result") (Schumacher, 2007). Using this type of matching will provide a lower number of matches because two records can match only on the basis of rules. This matching has difficulty accounting for the myriad of spelling and date-entry errors that may occur with a student record. A deterministic matching approach would return records with a 100% score based on the rules assigned. This method is best for smaller-sized data sets where the failure to match a record does not impact analysis (Schumacher, 2007). Although deterministic matching provides extremely accurate results, it is also less likely to return enough valid results for this study. As a result, I did not use it for this study.

Instead probabilistic matching was used. In this method, an algorithm assessed the likelihood that a match is correct and assigned a percentage "indicating the probability of a match" (Schumacher, 2007). This approach allowed for a larger number of addresses to be geocoded, although many had a probability score less than 100%. This was ideal because of the size of the dataset analyzed and because this method accommodated typing and spelling errors found in the dataset. Although used, the geocoded records output required a higher level of manual review to ensure the integrity of the data.

Once projected onto maps, spatial analysis also presents a problem related to population density. Projecting student enrollments may result in seemingly large clusters of enrolled students. A cluster is defined as occurring in a geographic distribution "when features are found in close proximity or when groups of features with similarly high or low values are found together (hot spots or cold spots)" (Mitchell, 2005, p. 148). The problem is that clusters are often the natural result of high or low population densities in a given block group.

A related issue is that representing enrolled students with dots can present issues related to the scale of the map. At a small scale where one inch represents hundreds of miles, a dot projected in two point type covers an area several miles wide. Because of this, the maps display the dots as layers *stacked* on top of one another. This stacking hides all dots in a given area, except for the dot on the top-most layer. As a result, a geographic location with 100 enrollees could appear as a single dot, the same as a location with only one dot. This results in the appearance of lower densities. The opposite effect occurs with larger scale maps. At a scale of 1:100,000 (used for the majority of maps in this study), a single dot in two point type represents an area with a diameter of nearly 150 feet. While all enrolled students are represented, the geographical space consumed by the individual student is much larger than the true space that would be occupied. This result is that a neighborhood could appear to have higher densities.

Conclusion

The purpose of this chapter was to provide an overview of the research methods employed in this dissertation. I have described the secondary data that will be used to answer my research questions. Also, the process of geocoding was presented, along with a discussion on census data block groups. I discussed the variables to be used, including those serving as proxies for the various forms of capital. The combination of census data and institutional data will enable this analysis to identify those pre-enrollment student characteristics influencing persistence, and allow this research to utilize measures that have not been used previously in attempting to understand. By identifying these student characteristics, enrollment managers may be able to utilize geographic-based variables associated with environmentally transferred capital as tools for strategically managing enrollments. Although limitations exist with the research, these findings will contribute to literature for persistence and graduation with respect to the use of spatial analysis within enrollment management.

CHAPTER FIVE: FINDINGS

This study investigated the relationship of student characteristics and their associated neighborhood characteristics to persistence and graduation outcomes at The University of Arizona. In this chapter, I begin by presenting institutional findings to provide greater context, and to assist the reader in understanding research outcomes. I then present the demographic and geodemographic information of students included in this study. I follow this with descriptive statistics for the dependent and independent variables used for analysis. This opening section is followed by three additional sections, presenting the findings of this research relating to the three main research questions.

Institutional Findings

Descriptive analysis.

To answer these questions as they relate to a single institution, this research was conducted at The University of Arizona, a major public research extensive university in the southwestern United States. This institution enrolled 27,379 undergraduates from all states and 125 countries during the final year of this four-year study (Fact book, 2007). The institution admits approximately 80% of all applicants, allowing for it to be classified as either a selective institution or as a moderately selective institution depending upon the classification criteria (OIRPS, 2008; Bowen, Chingos, & McPherson, 2009; U.S. News and World Report, 2009). The University of Arizona is considered the state's flagship research institution, and leads the state in overall quality of applicants, overall research dollars, and annual national rankings as reported by U.S. News & World Report (Fact book, 2007; 2009).

The time frame used in the selection of the sample, 2004-2007, corresponds to the first four enrollment years impacted by the university's creation of an office of enrollment management (Ota, 2002). The sample for this research included all domestic, full-time freshmen enrolled in the fall semesters from 2004 to 2007. As mentioned in the previous chapter, 21,838 student records were used for analyzing enrollments (Table 8).

Table 8
Enrollments by year

Category	2004	2005	2006	2007	Total
Enrollees	5043	5435	5396	5964	21838

Descriptive Statistics of Study Variables

In this section, I present the descriptive statistics of dependent and independent variables used in this study. These statistics are summarized in Table 9.

Table 9
Descriptive statistics of study variables

Variable name	Mean	SD
Retention year-one (RY1)	.80	.40
Retention year-four (RY4G)	.64	.48
Gender	.54	.50
Caucasian	.71	.45
Hispanic	.16	.37
Asian American	.07	.26
African American	.04	.18
Native American	.03	.16
In-state tuition (resident)	.66	.48
High school GPA	3.31	.48
Standardized test score	1112.53	161.89
Median family income	75588.95	35734
Percent of education, bachelors or higher	.45	.23
Percent of block group classified as urban area	.83	.35
Percent of block group classified as urban cluster	.07	.23.
Number of standardized tests	.1.95	.98
Applied X months before school	8.45	2.07
Distance from university	391.02	584.01

Descriptive statistics of the dependent variables.

Year-One retention.

The dependent variable used in this study is based on the students' enrollment in the year following their first year of attendance. A student beginning in the fall semester of 2004 was later checked for enrollment as of the 21st day of classes of the next

academic year. In this case, if the student was enrolled in classes in the fall semester of 2005, the variable for retention in year one (RY1) was set to 1. If the student was not enrolled, the variable was set to 0. The mean is .80 (SD = .40), indicating the institution's average year-one persistence rate for the four years of this study is 80%. Indeed, first-year persistence rates during the four years of analysis were stable, with all four years in this sample having an 80% year-one retention rate (Table 10).

Table 10
Retention year one, by year (percent in parenthesis)

Enrollment status	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
Persisted	4055	(80)	4351	(80)	4337	(80)	4749	(80)	17492	(80)
Did not persist	988	(20)	1084	(20)	1059	(20)	1215	(20)	4346	(20)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

Graduated/Year-Four retention.

For the 2004 cohort, 34% (N=1,705) of the 5,043 enrollees graduated at the end of four years. Since students still enrolled, but not yet graduated, would be considered persisting, I also included this group. This combination comprises the second dependent variable, RY4G—students who had either graduated, or remained enrolled after the fourth year of attendance (RY4G). In cases where persistence was indicated, the dependent variable RY4G was set to 1. In cases where the students were no longer enrolled, and had not graduated after the fourth year, RY4G was set to 0. The mean is

.64 (SD = .48), indicating the institution's graduate/persistence rate for the 2004 enrollees in this study is 63.6% (Table 11).

Table 11
Graduation rate, by year (percent in parenthesis)

Enrollment status	2004	(%)
Graduated or still enrolled	3206	(64)
Did not persist	1837	(36)
Total	5043	(100)

Note. The percentages may not add up to 100% due to rounding.

Descriptive statistics of the independent variables.

Gender.

As mentioned, literature has identified the potential influence of gender warranting its inclusion in this analysis. In Chapter Three, this is identified as a categorical variable. As such it is a dummy variable coded as 0 for male, and 1 for female. The mean therefore represents the percentage of enrollees with a gender classified as female. The ratio of female to male students remained stable during the four years of this study, with a mean indicating the university enrolled 54 females for every 46 males (SD = .5) (Table 12).

This is similar to national trends where females account for 54.7% of full-time enrollees compared to 45.3% for males (Freeman, 2004).

Table 12
Enrollments by gender/year (percent in parenthesis)

	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
Female	2747	(55)	2939	(54)	2907	(54)	3268	(55)	11861	(54)
Males	2296	(46)	2496	(46)	2489	(46)	2696	(45)	9977	(46)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

Race/Ethnicity.

A substantial volume of literature addressed race/ethnicity considerations in student persistence. This is a categorical variable having five possible classifications, African American, Asian American, Caucasian, Hispanic, and Native American. For selected analysis, dummy variables (binary distribution) were created for each with 1 indicating the enrollee was a member of the specific race and 0 indicating they were not identified by the specific race. The university study data evidenced a predominantly Caucasian undergraduate population at 71% (SD = .45), followed by Hispanic at 16%, (SD = .37), Asian American at 7% (SD = .26), African American at 4% (SD = .18), and Native American at 3% (SD = .16) (discrepancies due to rounding) (Table 13). From year one of the study to year four, major changes in racial composition of the class were not evident. This is comparable to the national enrollment trends during this time where 24% of enrollees at doctoral institutions were ethnic minority students (Wirt, Choy, Rooney, Hussar, Provasnik, & Hampden-Thompson, 2005). On average at universities throughout the U.S., African American students accounted for 12% of the student body, Asian Americans 6%, and Hispanics 10% (Wirt, Choy, Rooney, Hussar, Provasnik, &

Hampden-Thompson, 2005). This indicates The University of Arizona has a typical enrollment for Caucasian and Asian American students, but an atypically high percentage of Hispanic students, and a less than average percentage of African American students. Native American students were not referenced in the study.

Table 13
Enrollments by race/ethnicity/year *(percent in parenthesis)*

Race/ ethnicity	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
Caucasian	3626	(72)	3946	(73)	3735	(69)	4186	(71)	15493	(71)
Hispanic	779	(16)	799	(16)	923	(17)	946	(16)	3447	(16)
Asian American	360	(7)	379	(7)	413	(8)	399	(7)	1551	(7)
African American	161	(3)	183	(3)	174	(3)	245	(4)	763	(4)
Native American	117	(2)	128	(2)	151	(3)	188	(3)	584	(3)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

Residency status.

The student characteristic of residency is included due to its potential influence on persistence as identified by literature on social integration and affluence. This categorical variable was given a value of 1 for residents and 0 for nonresidents. During the time period covered in this study, the board of regents gave state universities greater flexibility in managing enrollments. Resulting policy changes allowed the university greater control over admissions criteria, directly impacting the geographical origination points of

enrolling students. Until these changes, the institution was forced by state policy to reject "nonresident applicants who were significantly more likely to succeed than some residents who actually enrolled, paying less than one quarter of the tuition offered by the student denied" (ABOR, 2005). As a result of the enrollment policy changes, a change in the class composition was present. Averaged over four years, resident enrollees comprised 66% of the incoming class (SD = .48) compared to 35% for nonresidents (due to rounding, this does not equal 100%) (Table 14). However, comparing year one of the study sample (2004) to year four (2007), the ratio of residents decreased from 69% of freshman enrollees to 64%, with nonresidents experiencing a reciprocal increase.

Table 14
Enrollments by residency/year (percent in parenthesis)

Residency	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
Resident	3464	(69)	3621	(66)	3417	(63)	3804	(64)	14306	(66)
Nonresident	1579	(31)	1814	(33)	1979	(37)	2160	(36)	7532	(35)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

High school grade point average.

With high school GPA identified by several past studies as the strongest indicator of college performance and persistence, it is included in this study (Alexander, Riordan, Fennessey, & Pallas, 1982; Mathiasen, 1984; Mouw & Khanna, 1993; Pike & Saupe, 2002; Astin, Tsui, & Avalos, 1996). GPA ranged from a low of 1.36 to 4.0 with a mean GPA 3.31 over the four years (SD = .48) (Table 15). Mean GPA for enrolling students remained constant when disaggregated by enrollment year.

Table 15
Distribution of High School GPA by enrollment year

GPA category	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
A or A+ (4.0)	335	(7)	350	(6)	292	(5)	298	(5)	1275	(6)
A- (3.7-3.99)	1141	(23)	1142	(21)	1066	(20)	1196	(20)	4545	(21)
B+ (3.3-3.69)	1311	(26)	1378	(25)	1429	(27)	1687	(28)	5805	(27)
B (3-3.29)	1007	(20)	1081	(20)	1085	(20)	1242	(21)	4415	(21)
B- (2.7-2.99)	691	(14)	870	(16)	759	(14)	946	(16)	3266	(14)
C+ (2.3-2.69)	458	(9)	507	(9)	604	(12)	542	(9)	2111	(10)
C (2.0-2.29)	85	(2)	96	(2)	140	(3)	48	(1)	369	(2)
<C (< 2.0)	15	(0)	11	(0)	21	(0)	5	(0)	52	(0)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Standardized tests (SAT/ACT test scores).

As mentioned, the influence of test scores on persistence and graduation is debatable (Mathiasen, 1984; Pike & Saupe, 2002; Vars & Bowen, 1998; Hezlett, et al., 2001; Bowen & Bok, 1998; Vars & Bowen, 1998). Scores from the SAT and ACT are included in this analysis, the mean test score is 1,113 (SD = 161.89) (Table 16). Average test scores during the study decreased from 1,119 in 2004 to 1,105 in 2007.

Table 16
Enrollments by standardized test score/year (percent in parenthesis)

Score range	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
400-499	0	(0)	1	(0)	1	(0)	0	(0)	2	(0)
500-599	4	(0)	1	(0)	2	(0)	5	(0)	12	(0)
600-699	26	(1)	21	(0)	23	(0)	24	(0)	94	(0)
700-799	98	(2)	105	(2)	119	(2)	131	(2)	453	(2)
800-899	309	(6)	327	(6)	358	(7)	375	(6)	1369	(6)
900-999	710	(14)	783	(14)	839	(16)	1006	(17)	3338	(15)
1000-1099	1015	(20)	1116	(21)	1199	(22)	1288	(22)	4618	(21)
1100-1199	1303	(26)	1380	(25)	1310	(24)	1450	(24)	5443	(25)
1200-1299	843	(17)	874	(16)	816	(15)	933	(16)	3466	(16)
1300-1399	521	(10)	564	(10)	508	(9)	534	(9)	2127	(10)
1400-1499	161	(3)	209	(4)	167	(3)	172	(3)	709	(3)
1500-1599	52	(1)	51	(1)	52	(1)	45	(1)	200	(1)
1600	1	(0)	3	(0)	2	(0)	1	(0)	7	(0)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

Median household income.

Median household income ranged from $2,499 to greater than $200,000, as the census does not disaggregate income levels above this level. The mean for this characteristic was $75,589 (SD = 35,734) (Table 17). Groups were divided into quintiles for analysis. All groups experienced in-category growth due to overall enrollment growth, with the notable exception of those in the lowest quintile. Those in

neighborhoods with median household incomes of $40,000 or below decreased by 9%. This group represented 17% of the first-time freshman population in 2004, but only accounted for 13% in 2007. The highest income quintile, those above $160,000, increased by 47%. However, due to a low number of enrollments in this group, they represented three percent of the first-time freshman population in 2004 and 2007. Other groups showed similar results, $40,000- $80,000 (in-category increase of 15%, overall decrease from 47% of first-time freshman to 46%); $80,000-120,000 (in-category increase of 31%, overall increase from 26% of first-time freshman to 29%); and $120,000-$160000 (in-category increase of 51%, overall increase from 7% of first-time freshman to 9%).

Table 17
Enrollments by median household income/year (percent in parenthesis)

Income range	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
0-40,000	865	(17)	812	(15)	742	(14)	789	(13)	3208	(15)
40,000-80,000	2386	(47)	2526	(46)	2436	(45)	2744	(46)	10092	(46)
80,000-120,000	1320	(26)	1554	(29)	1605	(30)	1724	(29)	6203	(28)
120,000-160,000	345	(7)	383	(7)	447	(8)	521	(9)	1696	(8)
>160,000	127	(3)	160	(3)	166	(3)	186	(3)	639	(3)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

First generation college student/family educational background.

Family educational background is measured by the proxy of the percent of males in the neighborhood with at least a bachelor's degree (education level). The sample ranges from a low of .00 to a high of 1. Mean for education level was .45 (SD = .23) (Table 18). I display the data as quintiles for consistency with the maps presented in Chapter Five. All five categories increased, with lowest to highest increasing in-category by 11, 14, 16, 27, and 43%. For all first-time freshman enrollments from 2004-2007, the lowest three educational attainment categories decreased by 1% each, and the highest increased respectively by 2 and 1%. A national report on education levels for 25- to 29-year-olds reported Caucasians were nearly twice as likely to have a bachelor's degree or higher when compared to African Americans and Hispanics in the same age group (Table 19)

Table 18
Enrollments by household education levels/year (percent in parenthesis)

% with bachelor's or higher	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
0-20%	929	(18)	963	(18)	937	(17)	1027	(18)	3856	(18)
20-40%	1284	(25)	1302	(24)	1267	(23)	1461	(24)	5314	(24)
40-60%	1403	(28)	1498	(28)	1512	(28)	1629	(28)	6042	(28)
60-80%	1182	(23)	1385	(25)	1398	(26)	1496	(25)	5461	(25)
>80%	245	(5)	287	(5)	282	(5)	351	(5)	1165	(5)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

Table 19
Percent with bachelor's degree or higher by race/ethnicity and gender (2000)

Race/ethnicity	Females	Males	Total
African American	17.4	18.4	17.8
Caucasian	35.8	32.3	34.0
Hispanic	11.0	8.3	9.7
All	34.0	27.9	34.0

Data from the National Center for Education Statistics (Wirt, Choy, Rooney, Hussar, Provasnik, & Hampden-Thompson, 2005)

Urban versus rural.

This variable is represented by two variables, for reasons mentioned in Chapter Four. The first variable, the percent of the neighborhood classified as an urban, had a mean of .83 (SD = .35). The second variable, percent of the neighborhood classified as an urban cluster, had a mean of mean of .07 (SD = .23). Whether or not a student resided in rural or urban neighborhood prior to enrollment may have important implications for retention research, specifically with regard to institutional fit. Specifically, a rural student may face two dissonance issues. The first results from the physical location The University of Arizona. The university resides in an urban area with a 2005 estimated metropolitan area population of nearly one million residents, and is classified as an urban area (U.S. Census Bureau, Population Division, 2006). Students from rural areas may have adjustment issues associated with this dramatic change from a rural lifestyle to the urban lifestyle. Secondly, the enrollee may have difficulty finding with peers at this institution where 76% of all enrollees originated from a neighborhood that was 100% urban, and 91% were from a neighborhood that was at least 50% urban.

Number of standardized tests taken.

This student characteristic is viewed as a proxy for social capital, in that the parent or other influencer may encourage the student to take multiple tests in order to achieve the best score possible. While I found no reference to the use of this variable within the literature, it serves as a proxy for social capital —an indicator of persistence based on the social capital literature discussed in Chapter Three. Although, not required by the admissions office, all participants in this sample submitted standardized test scores. Number of tests ranged from 1 to 12 tests with a mean of 1.95 (SD = .98) (Table 20). Rank was in order from the lowest number of tests to the highest— one test (38.57%), two tests (37.26%), three tests (17.41%), four tests (5.13%) and greater than four tests (1.63%).

Table 20
Enrollments by number of tests scores submitted/year (percent in parenthesis)

Number	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
1	1844	(37)	1997	(37)	2107	(39)	2475	(41)	8423	(39)
2	1969	(39)	2066	(38)	1942	(36)	2160	(36)	8137	(37)
3	882	(17)	1003	(18)	937	(17)	980	(16)	3802	(17)
4	277	(5)	267	(5)	295	(5)	280	(5)	1119	(5)
5	49	(1)	79	(1)	85	(1)	52	(1)	265	(1)
6	18	(0)	15	(0)	21	(0)	13	(0)	67	(0)
7	4	(0)	4	(0)	5	(0)	2	(0)	15	(0)
8	0	(0)	4	(0)	3	(0)	0	(0)	7	(0)
9	0	(0)	0	(0)	0	(0)	0	(0)	0	(0)
10	0	(0)	0	(0)	1	(0)	0	(0)	1	(0)
11	0	(0)	0	(0)	0	(0)	1	(0)	1	(0)
12	0	(0)	0	(0)	0	(0)	1	(0)	1	(0)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

Applied date in number of months before school.

This student characteristic also is viewed as a proxy for social capital, in that the parent or other influence may encourage the student to apply early, again based on the social capital literature discussed in Chapter Three. Early application would allow the student to weigh options and make an informed decision to attend an institution fitting the student's individual needs. Again, I found no reference to the use of this variable within the literature, but it serves as a proxy for social capital. The range is from 14 months prior to the start of school to zero months (Table 21). The mean was 8.45 (SD = 2.07).

Typically, the published application deadline was four to five months before the mid-August start of the new school year; however, application deadlines were extended during the study period.

Table 21
Enrollments by application submission in months before start of school/year (percent in parenthesis)

Months before start of school	2004	(%)	2005	(%)	2006	(%)	2007	(%)	Total	(%)
0	0	(0)	9	(0)	2	(0)	2	(0)	13	(0)
1	8	(0)	15	(0)	8	(0)	9	(0)	40	(0)
2	21	(0)	25	(0)	34	(1)	15	(0)	95	(0)
3	51	(1)	64	(1)	64	(1)	61	(1)	240	(1)
4	117	(2)	173	(3)	303	(6)	215	(4)	808	(4)
5	185	(4)	217	(4)	297	(6)	261	(4)	960	(4)
6	374	(7)	380	(7)	580	(11)	547	(9)	1881	(9)
7	367	(7)	547	(10)	560	(10)	505	(8)	1979	(9)
8	575	(11)	726	(13)	1056	(20)	1145	(19)	3502	(16)
9	1028	(20)	831	(15)	1077	(20)	1358	(23)	4294	(20)
10	1552	(31)	1056	(19)	1324	(25)	1105	(19)	5037	(23)
11	763	(15)	1315	(24)	29	(1)	575	(10)	2682	(12)
12	2	(0)	46	(0)	43	(1)	145	(2)	236	(1)
13	0	(0)	12	(0)	8	(0)	12	(0)	32	(0)
14 or >	0	(0)	19	(0)	11	(0)	9	(0)	39	(0)
Total	5043	(100)	5435	(100)	5396	(100)	5964	(100)	21838	(100)

Note. The percentages may not add up to 100% due to rounding.

Summary of Descriptive Statistics

In this section, I established the foundation for further study into the relationship of student characteristics, associated neighborhood characteristics, year-one persistence and graduation outcomes at The University of Arizona. By presenting institutional findings, along with demographic and geodemographic information of students I provided greater context for the upcoming chapters on spatial distributions of enrollments, student characteristics predicting year-one persistence, and student characteristics predicting graduation. This section's presentation of descriptive statistics for the dependent and independent variables has provided a foundation that I build upon as I move on to the next three sections detailing the findings of this research.

Research Question 1: The Spatial Patterns of Enrollments

To answer the first part of research question one—what are the spatial patterns of enrollment with respect to resident and nonresident enrollees— I project enrollee data onto maps of the United States at various scales. As illustrated in Figure 2 (in Chapter Four), significant stacking of enrollees (represented by small circles) occurs when presenting information at a scale representing the entire continental United States. As mentioned, stacking as used in this dissertation refers to the overlap of the dots representing enrolled students. Although visually deceiving, the purpose of presenting maps at such a scale is to identify general enrollment patterns—most notably the institution's national reach. While the presentation of maps begins with a relatively small scale map, the maps increase in scale allowing specific demographic and geographic details to emerge.

It is also important when reading the following maps to understand that the display of maps in black and white poses a problem related to the use of gradients as representations of data categories. The more categories used, the greater detail presented; however, it becomes more difficult to visually delineate the breaks when using gradients of black and white (Monmonier, 1996). For all maps with categories, I present the data in quintiles based on examples from Edward Tufte's *The Visual Display of Quantitative Information* (2001). Defining the categories for these breaks may introduce bias into the research, and distort the findings (Monmonier, 1996). For this research, categorical cutoffs for each quintile were selected based on the context of the data. Several of the choices for categories are explained further in the description of individual maps.

Spatial Patterns of enrollments for resident and nonresident enrollees.

I begin this chapter with Figures 3 and 4 to focus on enrollments in the western and eastern halves of the US, providing greater detail than Figure 2. In the first map (Figure 3), it is evident that the institution has higher concentrations of enrollees closer to the university's physical location in Tucson, Arizona. Not surprisingly, the most populous state in the nation, California, provided 2,922 enrollees, far more than any state other than Arizona (Table 22). In spite of their proximity to the university, less populous states bordering Arizona did not provide large numbers of enrollees: New Mexico (203), Nevada (197), and Utah (51). After California, Texas (431) provides the greatest number of enrollees, although it is more difficult to see on the maps because of the amount of area that Texas geographically covers, and because the state's enrollments are divided onto both Figure 3 and Figure 4.

Table 22
Top 10 states for enrollment 2004-2007, western extent of the United States

State	Enrollments
Arizona	14483
California	2922
Texas	431
Washington	355
Colorado	296
New Mexico	203
Nevada	197
Oregon	171
Utah	51
Idaho	45

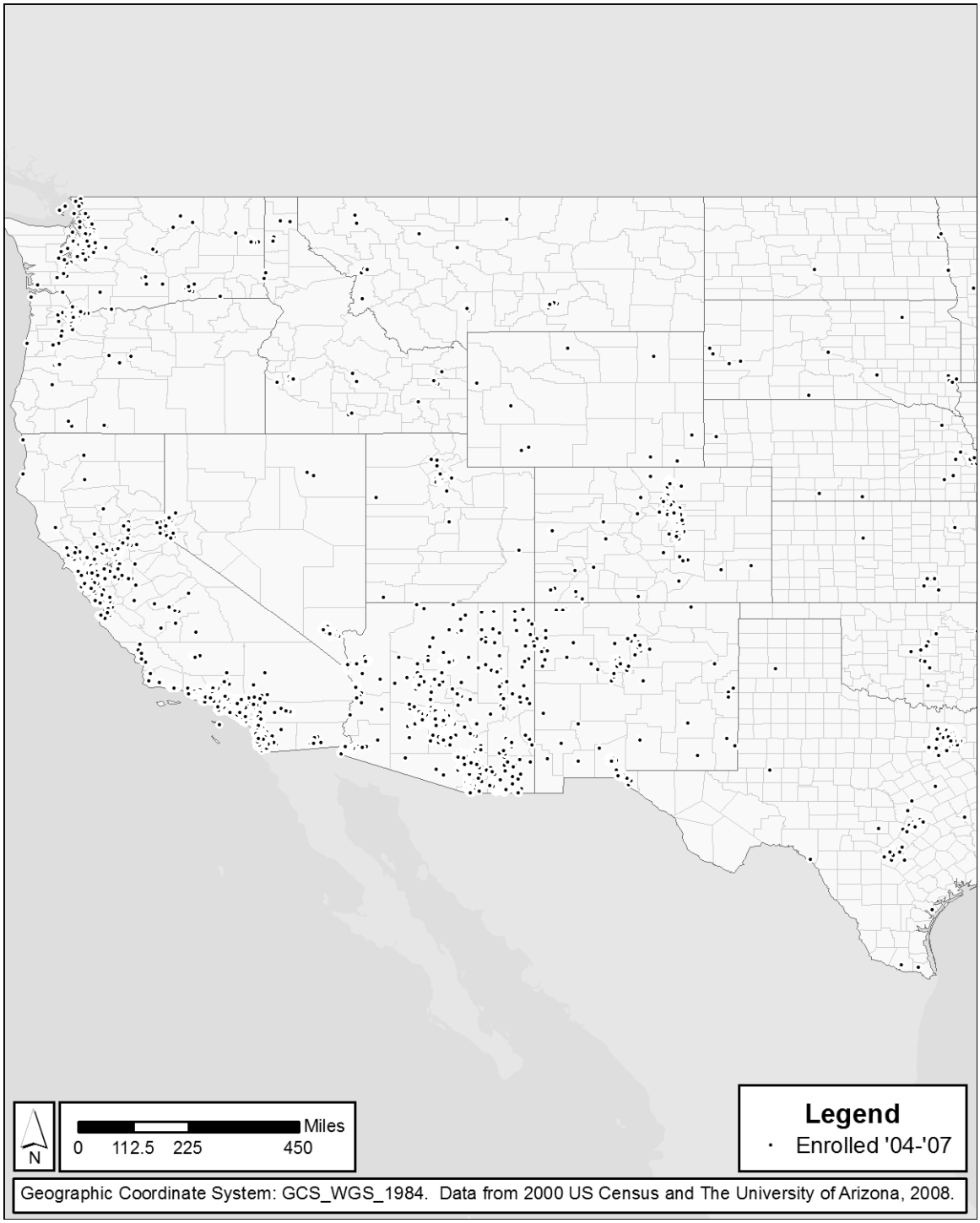

Figure 3. Enrollments for the lower 48 states, western extent. At this reduction, significant stacking still occurs for enrollees displayed.

In Figure 4, it becomes easy to identify enrollments concentrated in selected Midwestern and Eastern states and regions. As visually apparent in this map, the majority of enrollments in the eastern portion of the U.S. were from Illinois (490), New York (290), and New Jersey (201) (Table 23).

Table 23
Top 10 states for enrollment 2004-2007, eastern extent of the United States

State	Enrollments
Illinois	490
New York	290
New Jersey	201
Pennsylvania	155
Massachusetts	143
Minnesota	139
Ohio	139
Connecticut	97
Maryland	92
Michigan	89

Figure 4. Enrollments for the lower 48 states, eastern extent. At this reduction, significant stacking still occurs for enrollees displayed.

In Figure 4, it was difficult to distinguish the points of origination for California's 2,922 enrollees. They appear to be from coastal areas, but beyond that specific areas are difficult to identify. In Figure 5, it becomes easier to distinguish as I present enrollments from California, beginning at San Francisco in the north, down to San Diego in the south. Along with Sacramento and the Los Angeles metropolitan area, these regions provide the greatest number of California enrollees for the institution. It appears that the area referred to as Inland California generates few enrollments, as does the coastal area between San Francisco and Los Angeles.

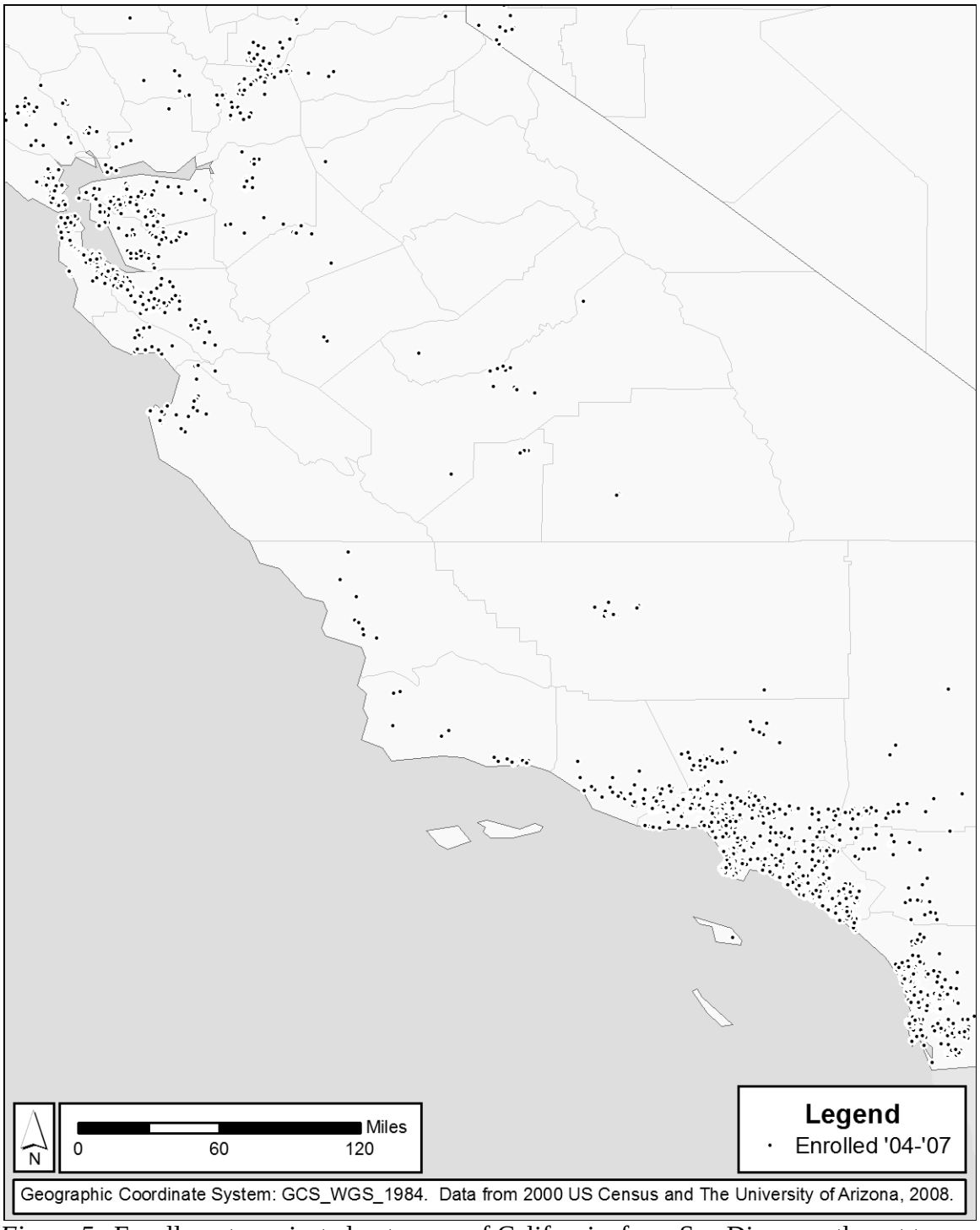

Figure 5. Enrollments projected onto map of California, from San Diego northwest to San Francisco.

Figure 6 presents a map showing enrollments for Arizona. As the map scales become larger, nuances of the enrollment patterns emerge. While Figure 4 gave the visual implication that resident enrollments were distributed throughout the state, Figure 6 illustrates that resident enrollees originate mainly from the state's two major population centers, Tucson and Phoenix. The Navajo and Hopi reservations in northeastern section of the state are sparsely populated, as are the western deserts, and the Grand Canyon region in the north—all yielding few enrollees. The map is less a representation of state areas where the university recruits well, and more a representation of the state's population distribution.

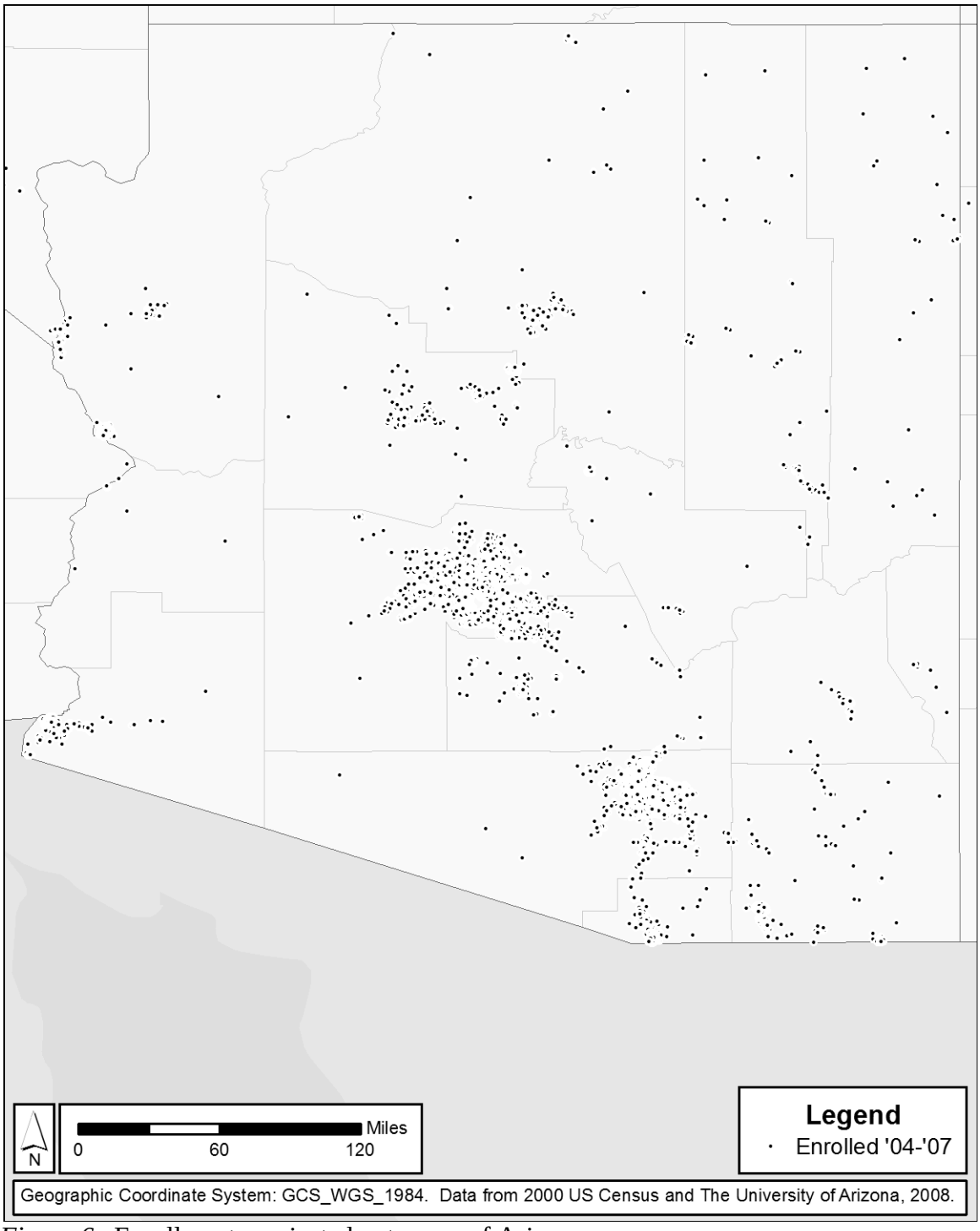

Figure 6. Enrollments projected onto map of Arizona.

These initial maps provide the general context for this analysis, but do not provide detailed information on the spatial distribution of enrollments. To answer this, I next present localized results of the enrollment mapping. In Figures 7-11, spatial findings are presented at the block group level. From these maps, I illustrate how enrollments vary across block groups. These maps allow for visualization at a higher level of detail, and are used throughout to show distributions of enrollees, by income levels, educational attainment and percent of minorities. I chose to present these detailed maps at a 1:100,000 scale to allow for an area small enough to show detail, but large enough to include a significant number of enrollees. Enrollments for four areas are presented. Since the majority of enrollees are from the Tucson metropolitan area, I present these findings in two maps, Figure 7 and Figure 8. In Figure 9, I present Phoenix, the university's second largest market (post-study, this became the largest market). Figure 10 presents Nogales, Arizona, a smaller city near the international border with Mexico. This city's population is predominantly Hispanic and was selected to provide insight into ethnic minority communities with respect to enrollment at The University of Arizona. In Figure 11, I use a map of Calabasas, California, a community in the western portion of the Los Angeles metropolitan area, to illustrate differences between resident and nonresident neighborhoods. Density of enrollees in nonresident neighborhoods is more difficult to display at the 1:100,000 scale because they produce few enrollees. Calabasas was selected because it yielded the highest number of nonresident enrollees in this sample.

Spatial distributions of enrollment for Tucson

Figures 7-8 show the spatial distributions of enrollments for Tucson. The Tucson metropolitan area had an estimated population of nearly 950,000 residents in 2005 (U.S. Census Bureau, 2007b). Geographical impediments to population distribution include the Catalina Mountains on the northwest portion of the northern extent, the airport, air force base, and sparsely populated Tohono O'odham reservation on the southern extent. The downtown area with few residential dwellings also appears to influence distributions. The Tucson Mountain foothills influence population distribution on the western section of both maps.

From these two maps, it appears enrollments from Tucson are more concentrated in the northern and central (near the university) portions of the city. In spite of geographical limitations to population density in the Catalina Foothills area, a large number of enrollees have addresses in this area. Although densities differ, enrollments are distributed throughout the entire metropolitan area.

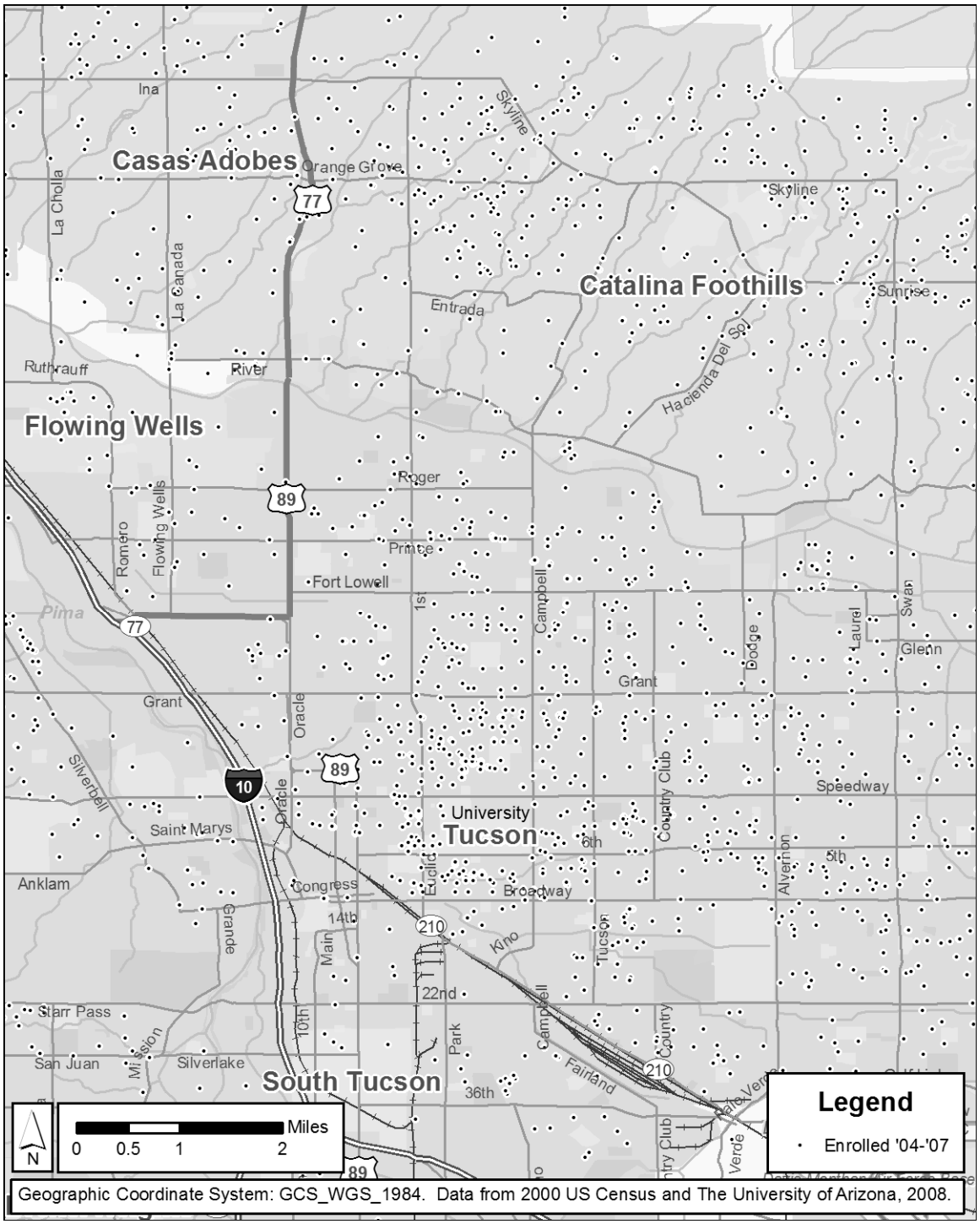

Figure 7. Enrollments projected onto map of the northern extent of Tucson, Arizona.

Figure 8. Enrollments projected onto map of the southern extent of Tucson, Arizona.

Spatial distributions of enrollment for Phoenix

The Phoenix metropolitan area had an estimated population of four million, in a state where the total estimated population was six million residents (U.S. Census Bureau, 2007b). Like many western cities, the population is spread out compared to eastern cities. The Phoenix metropolitan area covers 475 square miles (U.S. Census Bureau, 2007b). Figure 9 presents the central and southern portion of Phoenix. Limitations to residential population distribution include parks, mountain preserves, the airport and the downtown area. Most enrollments on this map are from the northwestern portion of the city of Phoenix (not metro area), Paradise Valley and Scottsdale.

Examining the map indicates areas of enrollments, and non-enrollments that are not related to geographical features. These areas appear to be the result of demographic features, rather than geographic features, as shown in subsequent maps.

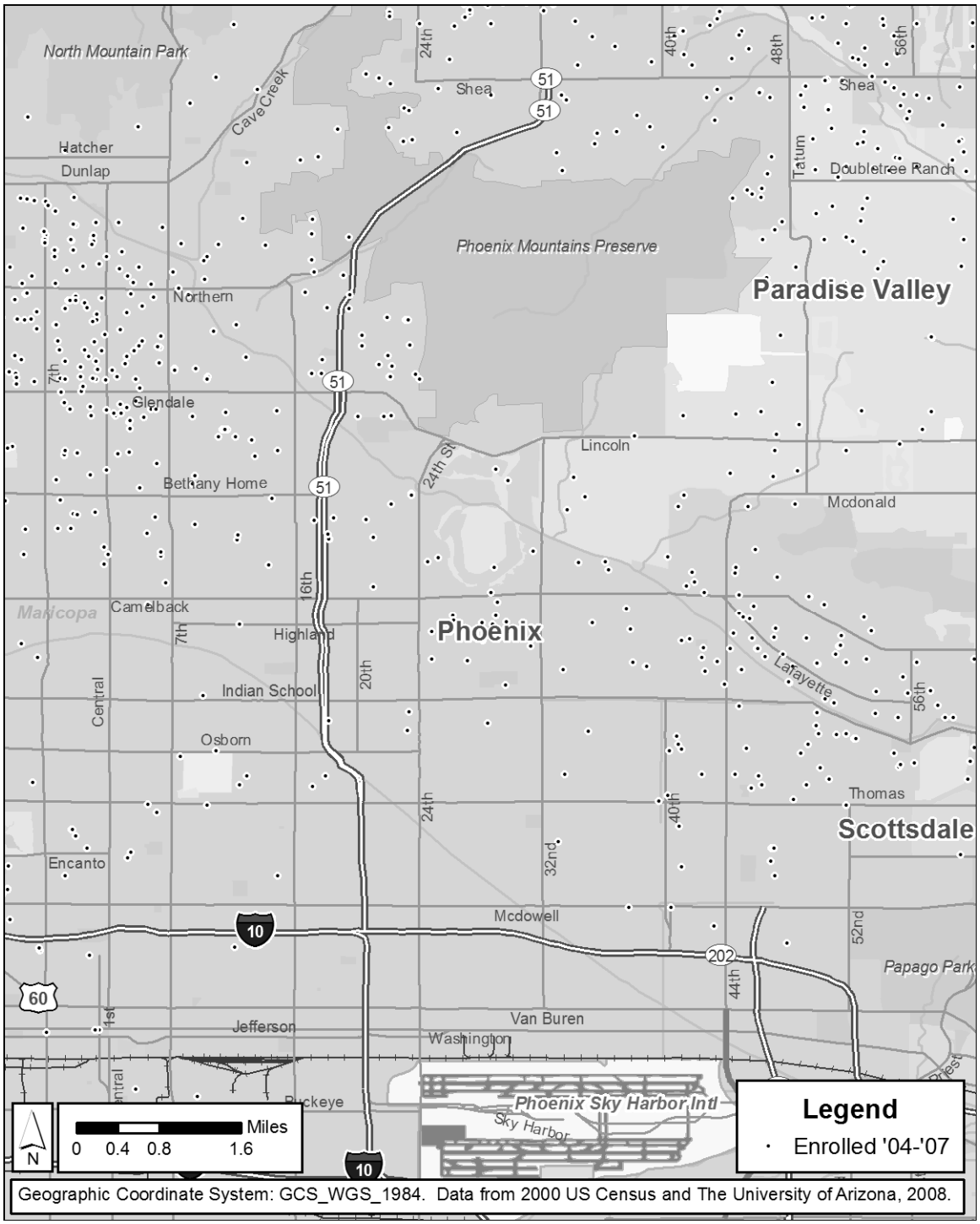

Figure 9. Enrollments projected onto map of Phoenix, Arizona.

Spatial distributions of enrollment for Nogales.

The city of Nogales has a population of approximately 21,000 residents. Located on the international border with Mexico, the city is 68 miles south of the University of Arizona. Although few enrollees are shown on the map, Nogales serves the purpose of illustrating enrollment from a predominantly Hispanic city, and surrounding area. As illustrated in Figure 10, most enrollments come from the city, with the rural areas surrounding Nogales providing a few enrollees.

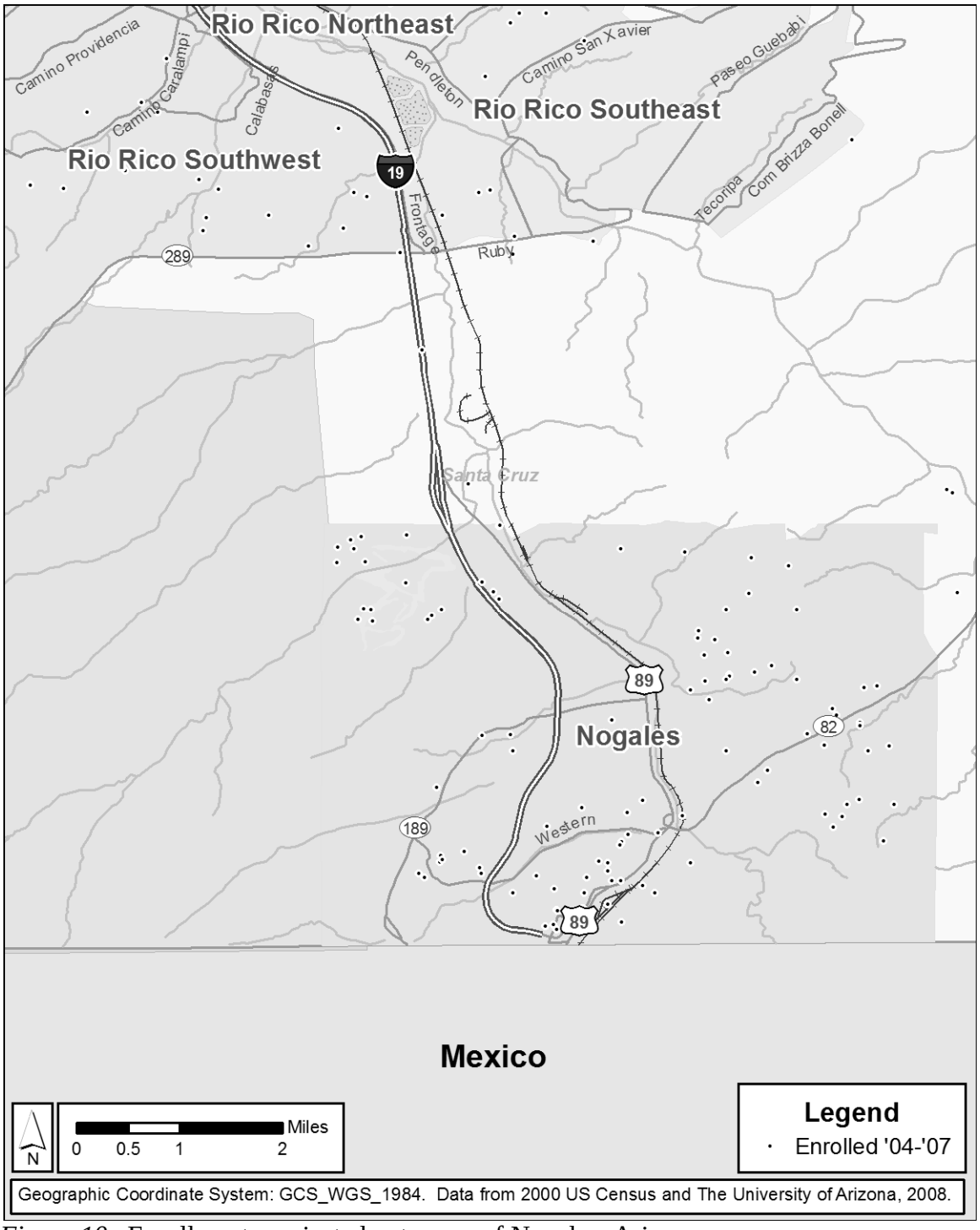

Figure 10. Enrollments projected onto map of Nogales, Arizona.

Spatial distributions of enrollment near Los Angeles.

The Los Angeles metropolitan area has a population of approximately 13 million residents (U.S. Census Bureau, 2007b). The map area is focused on the western suburbs of Los Angeles, in an area providing the greatest number of nonresident enrollees in the sample. Although few details can be gleaned from Figure 11, it appears enrollments are concentrated in specific neighborhoods, most notably near Calabasas. Areas in the northeast portion of the map, yield few enrollees compared to the other portions of the map. Obviously, shifting the area featured in the map would alter the way this appears, and the results of plotted distributions. Again, the reason for selection of this neighborhood was due to the high number of enrollees from the Calabasas neighborhood. As an enrollment management perspective, showing areas with high enrollments, next to areas with low enrollments is desirable to elucidate the neighborhood demographics underlying enrollment data.

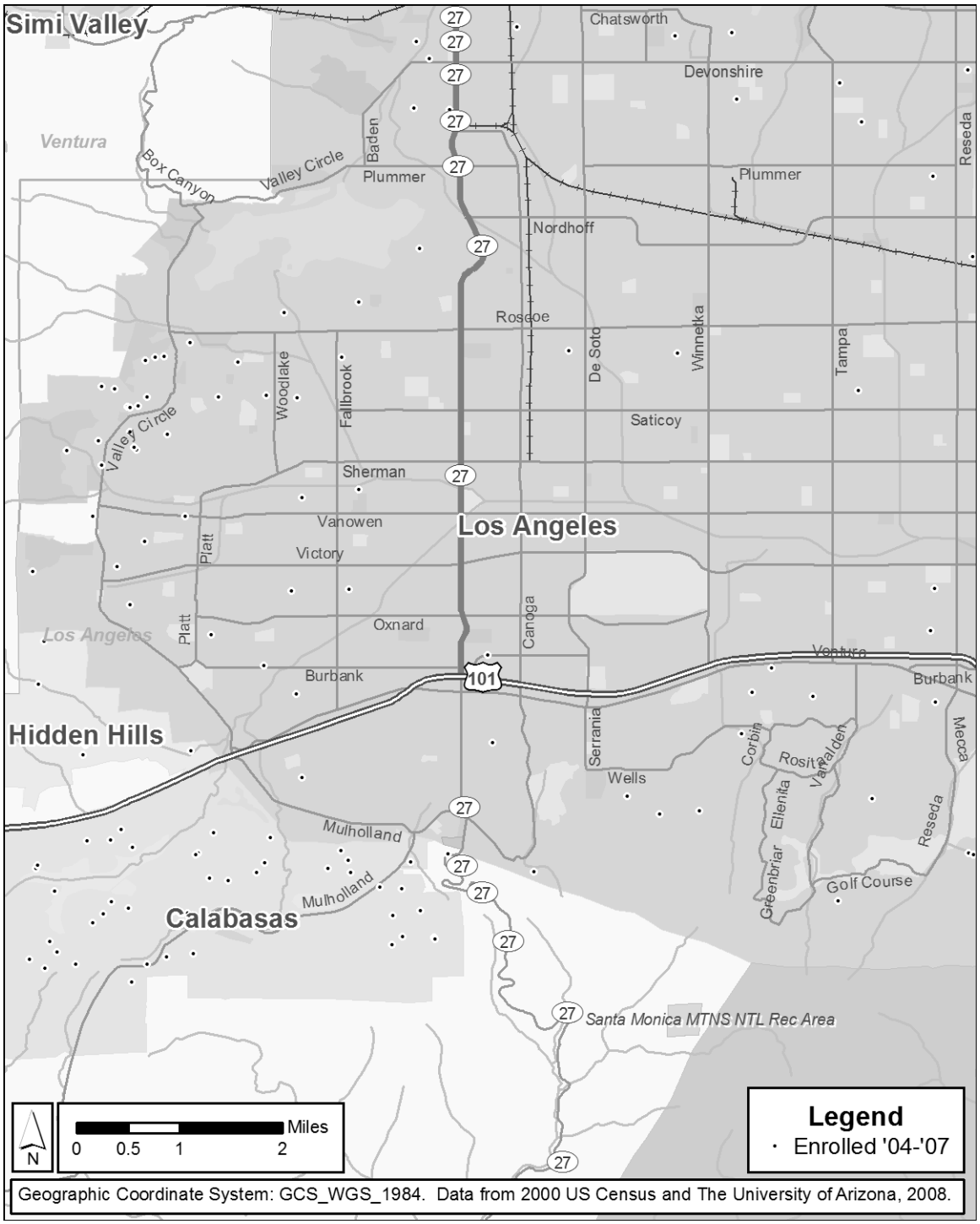

Figure 11. Enrollments projected onto map of western extent of Los Angeles, California.

Spatial patterns of enrollments by median household income levels.

With a greater understanding of enrollments by geography as presented in Figures 7-11, I now present findings on the spatial distributions of enrollments with respect to neighborhood household income levels. For the sample, the mean household income was 75,559 (SD=35,734). In Figures 12-16, Median household income levels from the census are disaggregated in quintiles. Using natural breaks at each $40,000 increment in the range contained in the census, the mid-level displayed has a starting point approximately $5,000 above the sample's mean. I maintain the layer featuring enrollments, as represented by dots on the map.

Spatial distributions of median household income levels for Tucson.

For Tucson enrollments (Figures 12-13), enrollments appear to be distributed in neighborhoods, with little difference among income levels. In Figure 12, even the lowest income neighborhoods are represented by a substantial number of enrollees. In Figure 13 representing the southern extent of Tucson, the lowest quintile appears to yield fewer enrollees.

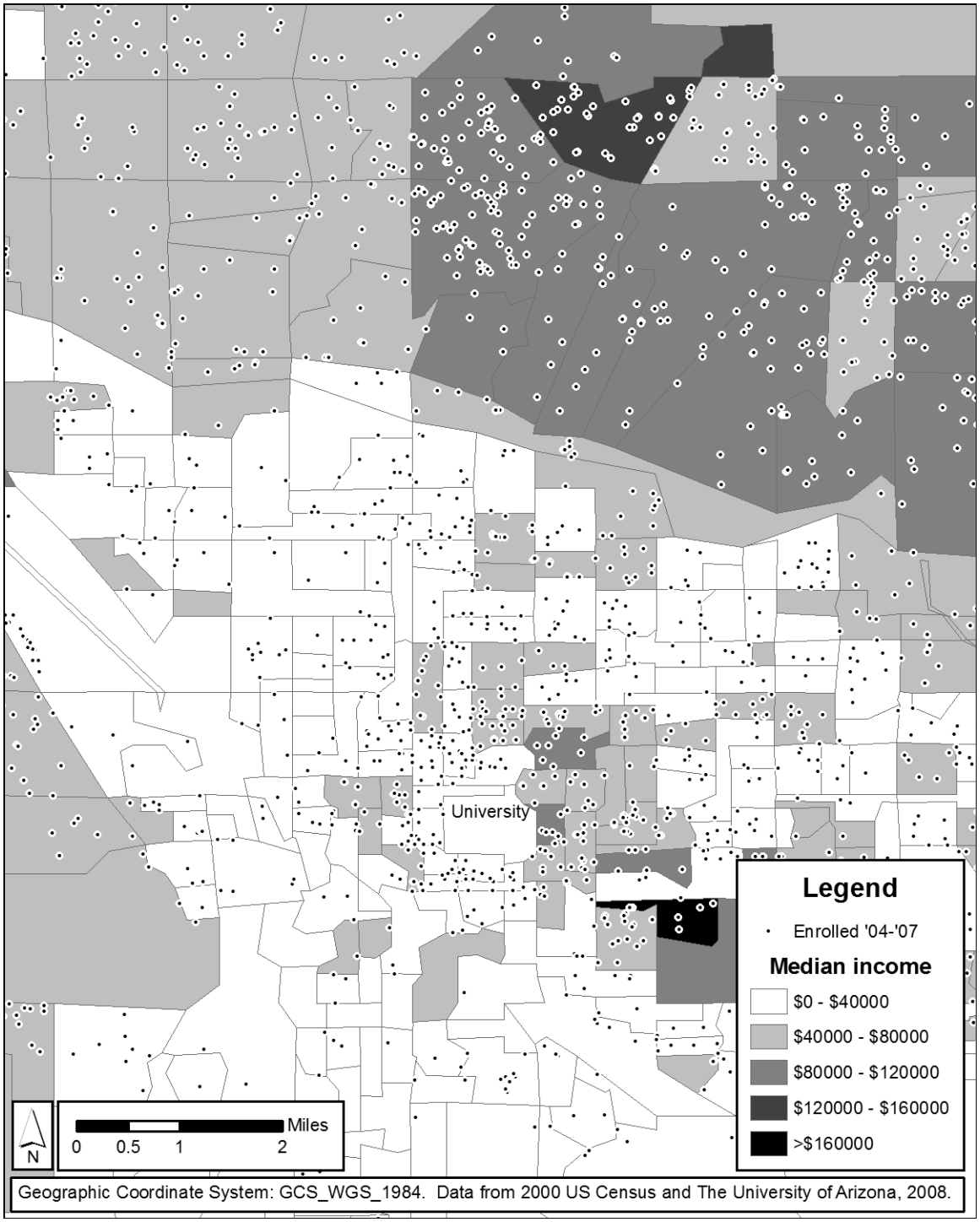

Figure 12. Enrollments with median income levels for block groups for the northern extent of Tucson, Arizona.

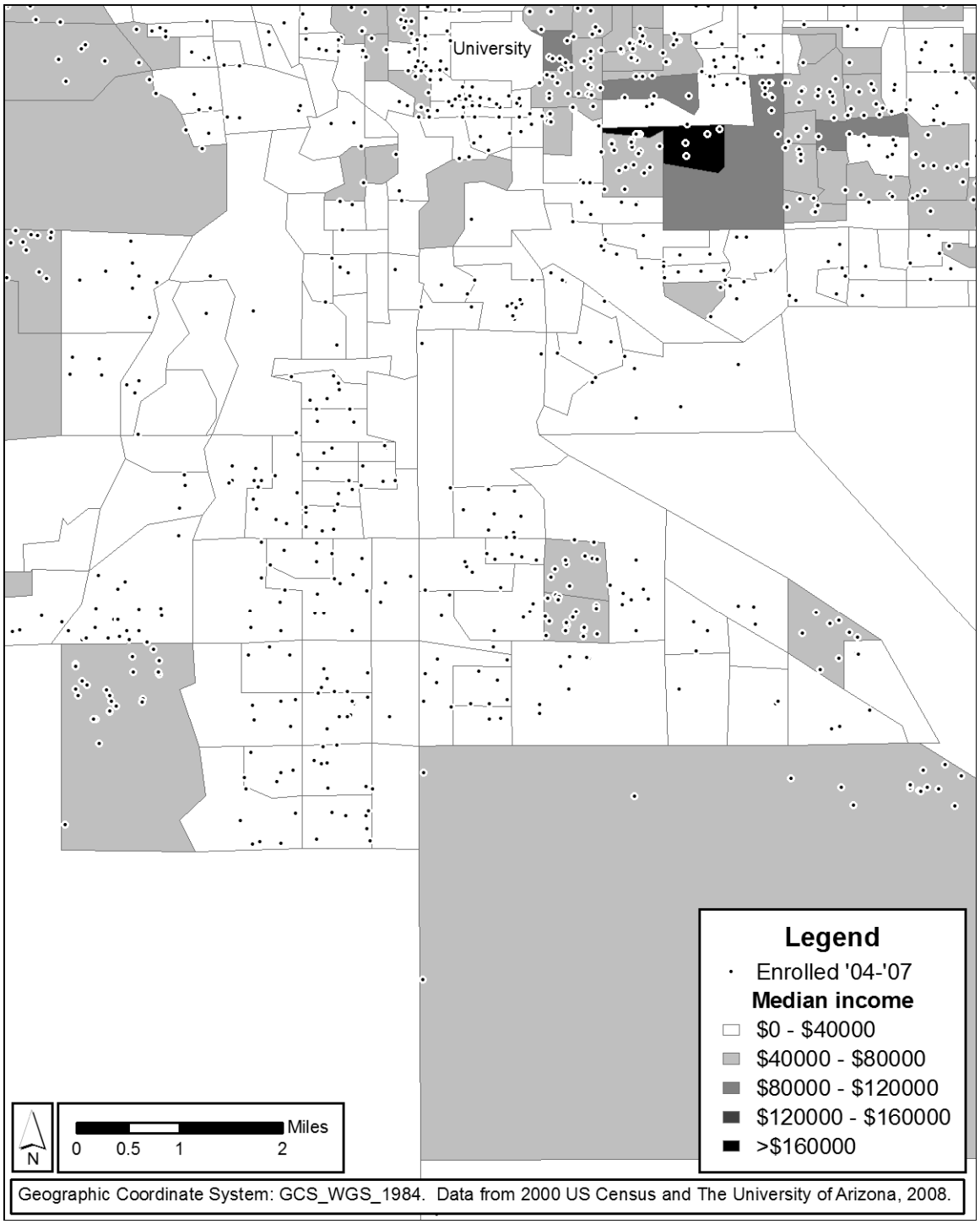

Figure 13. Enrollments with median income levels for block groups for the southern extent of Tucson, Arizona.

Spatial distributions of median household income levels for Phoenix.

Unlike the Tucson map, the map of Phoenix (Figure 14) begins to illustrate a definite difference in enrollments from the different neighborhood income levels. In Phoenix few enrollees from low-income neighborhoods are evident, and high concentrations of enrollees can be observed from the higher income areas. The large high income neighborhood in the northern part of the map with no enrollees is a mountain with few residents—although those few appear to be wealthy.

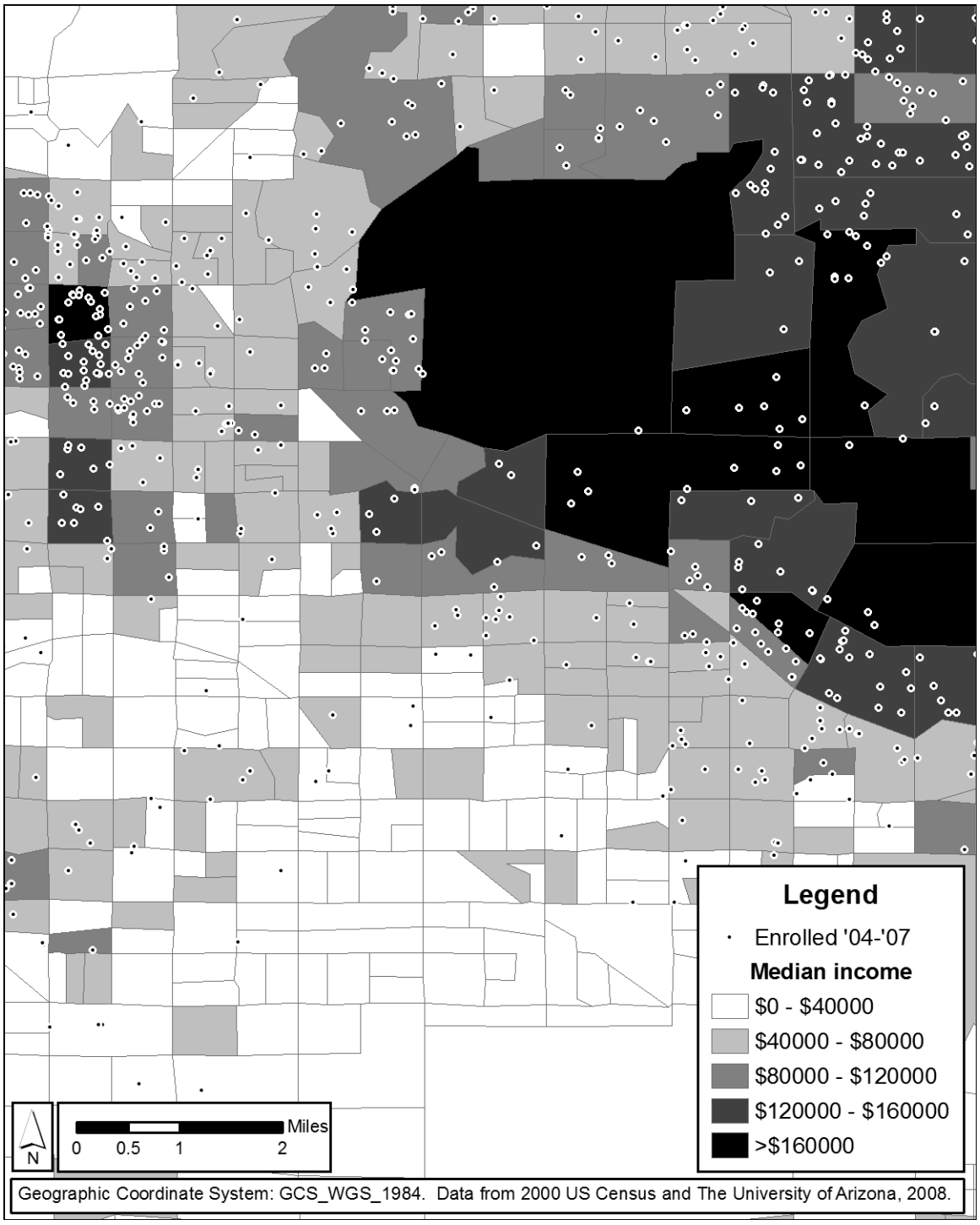

Figure 14. Enrollments with median income levels for block groups for Phoenix, Arizona.

Spatial distributions of median household income levels for Nogales.

For Nogales enrollments (Figure 15), the enrollment pattern appears to be more similar to Tucson with a substantial number of enrollees coming from the lowest income neighborhoods. Outlying rural areas along the eastern edge of the map are less densely populated. The apparent progression of dots from south to north follows the route of the interstate.

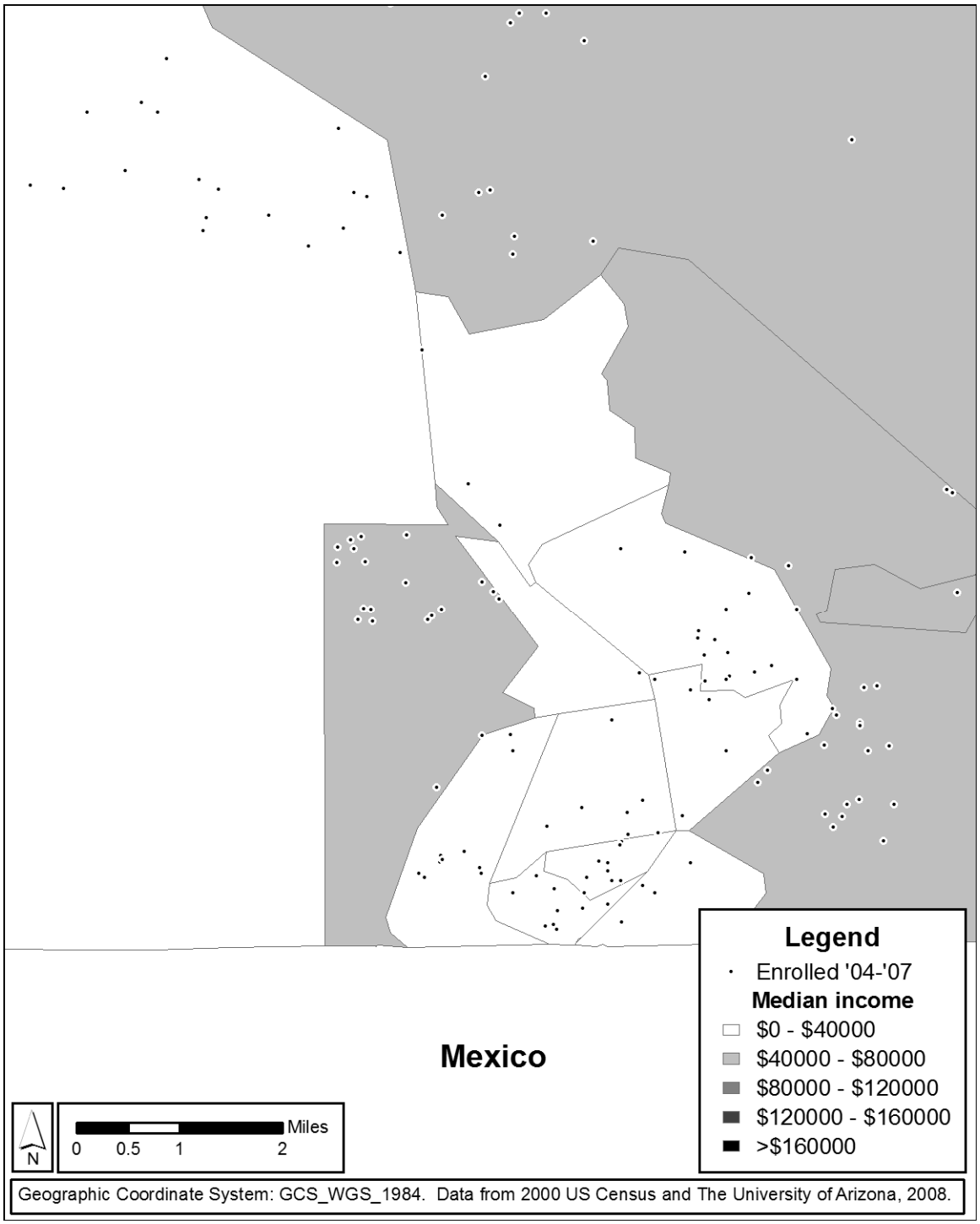

Figure 15. Enrollments with median income levels for block groups for extent of Nogales, Arizona.

Spatial distributions of median household income levels near Los Angeles.

For the Los Angeles enrollments (Figure 16), the enrollment patterns once again shift to higher income neighborhoods, this time skipping even the median income neighborhoods. Enrollments are concentrated in those neighborhoods with median income levels above $80,000 with the majority in neighborhoods above $120,000. As mentioned earlier, these nonresident students paid $9,000 to $11,000 more in tuition each year of this study.

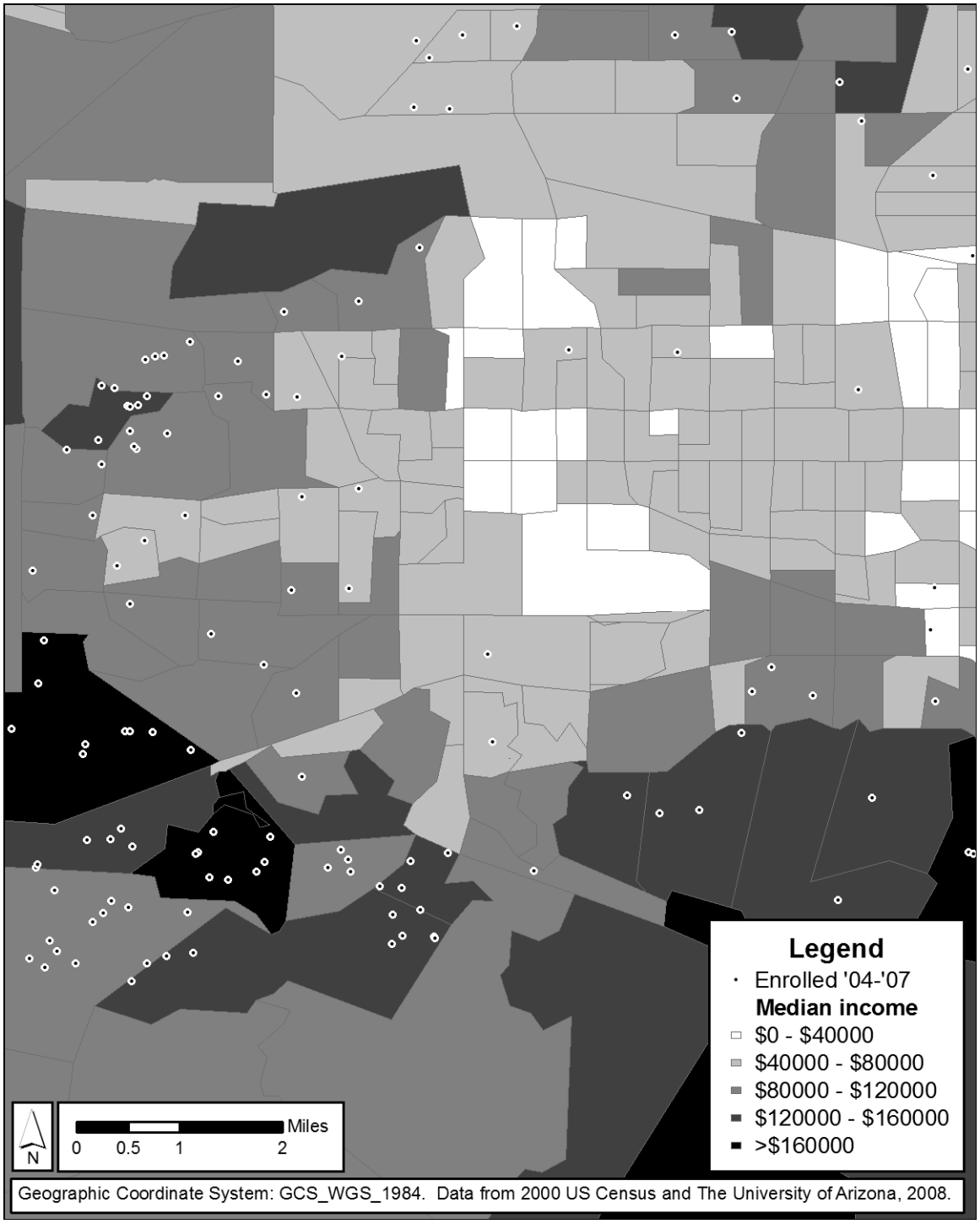

Figure 16. Enrollments with median income levels for block groups for western extent of Los Angeles, California.

Spatial patterns of enrollments by neighborhood education levels.

While not a definitive indicator of whether a student is a first generation student, the ability to create a map with neighborhood education levels enables a greater understanding of the background of students enrolling at The University of Arizona. I now present findings on the spatial distributions of enrollments with respect to neighborhood education levels. For this sample the percent of a neighborhood with a bachelor's degree or higher had a mean of .45 (SD=.23). In Figures 17-21, neighborhood education levels from the census are disaggregated in quintiles. Using natural breaks at each 20% increment in the range contained in the census, the mid-level (41-60%) incorporates the mean. From viewing the following maps, enrollments appear to be related more to education levels than income levels when contrasting the associated maps.

Spatial distributions of neighborhood education levels for Tucson

Levels of education in the Tucson metropolitan area are more sharply divided than income levels as illustrated in Figures 17-18. Neighborhoods with higher education levels are concentrated in the north, and near the university. While the northern distribution appears evenly distributed as far as education levels, the southern extent stands in sharp contrast. The southern portion of the city has a large swath of neighborhoods with the 0-20% categorization for the level of education.

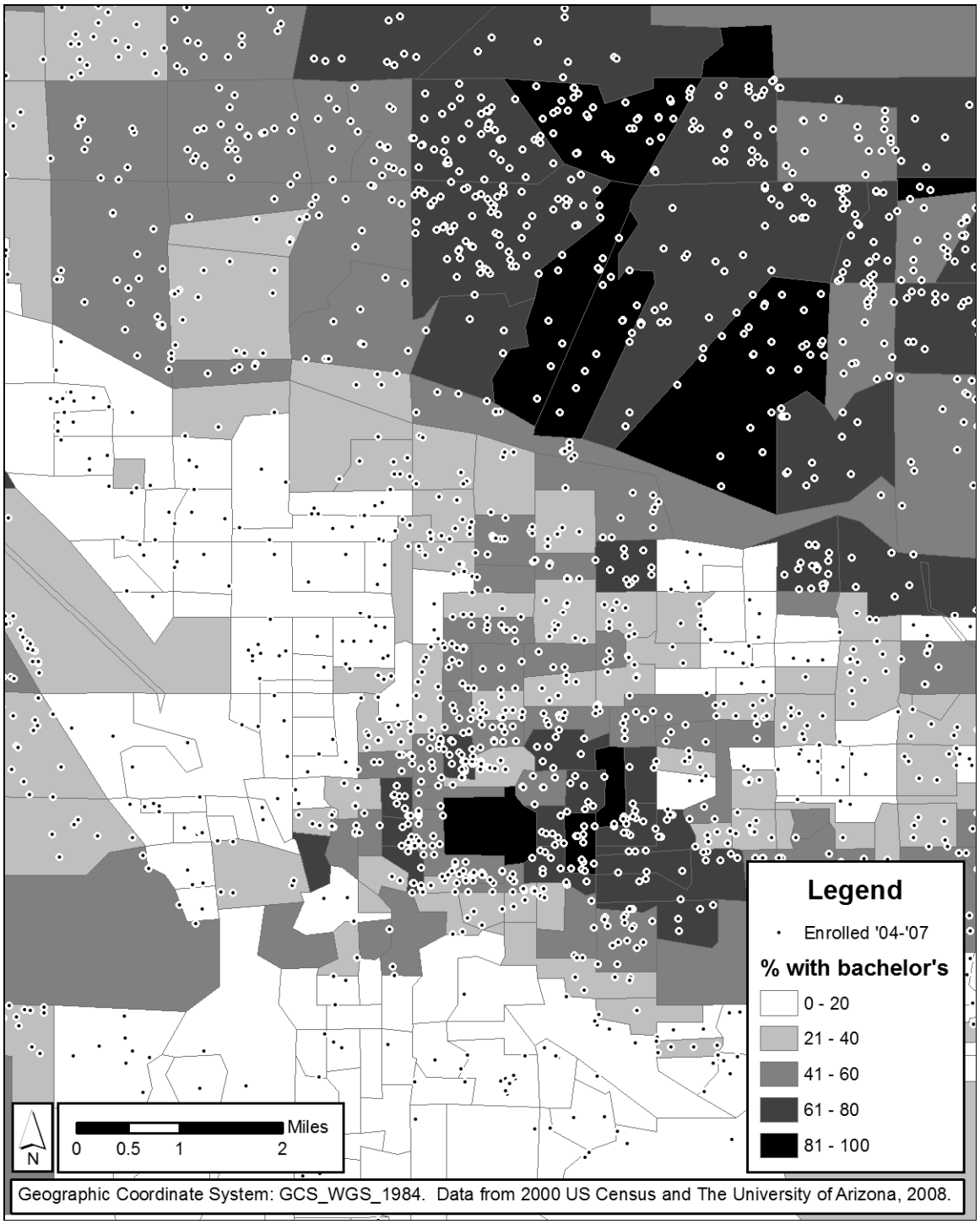

Figure 17. Enrollments with education levels as a percent (bachelor's degree or higher) for block groups for the northern extent of Tucson, Arizona.

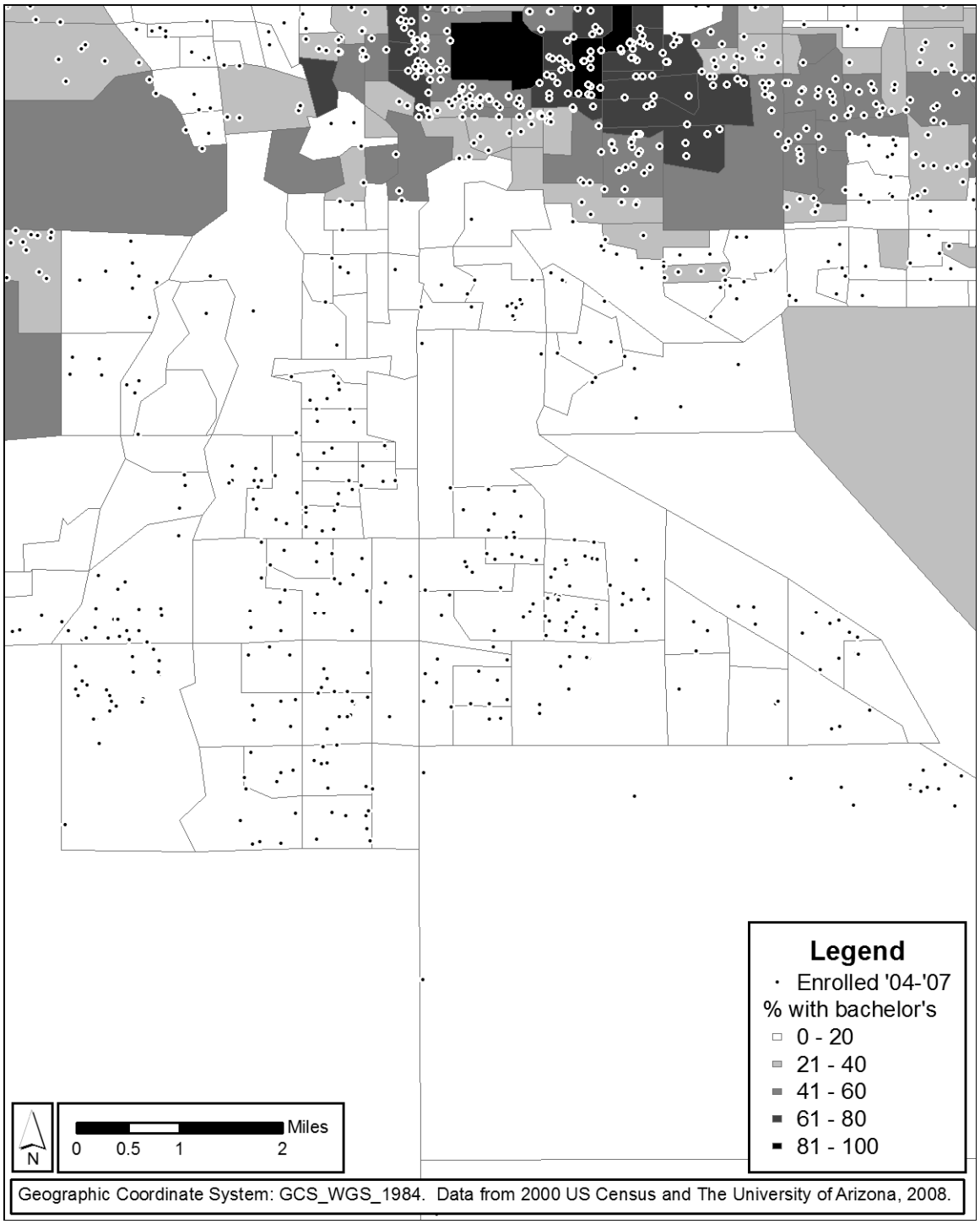

Figure 18. Enrollments with education levels as a percent (bachelor's degree or higher) for block groups for the southern extent of Tucson, Arizona.

Spatial distributions of neighborhood education levels for Phoenix

Once again, levels of education appear to be concentrated, more than with other neighborhood characteristics. Here, the southern portion of the map has low levels of education while the northeast section has higher than average levels. Again, unlike Tucson, enrollments in Phoenix seem to follow educational levels, concentrating in the areas with higher education levels.

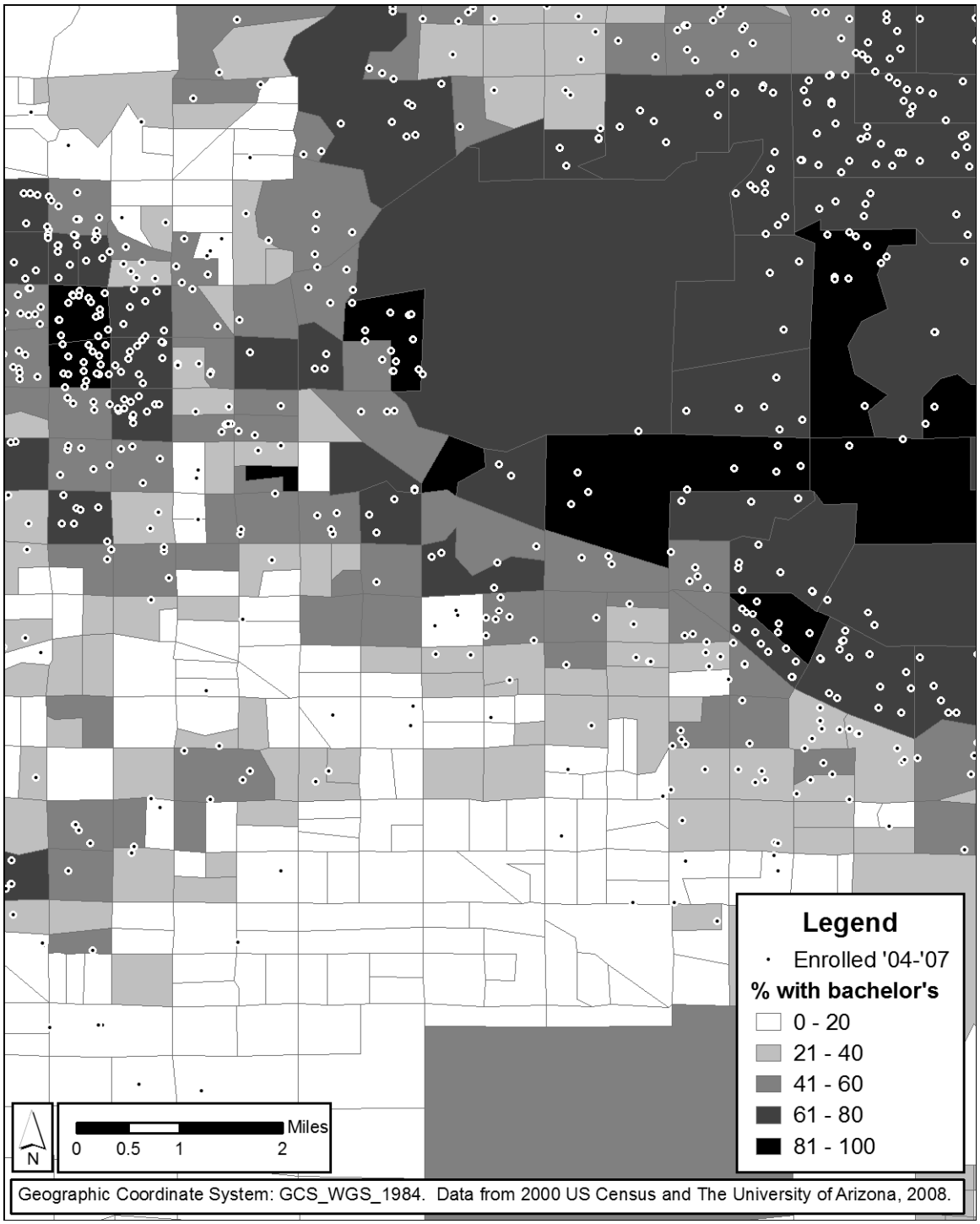

Figure 19. Enrollments with education levels as a percent (bachelor's degree or higher) for block groups for the southern extent of Phoenix, Arizona.

Spatial distributions of neighborhood education levels for Nogales

Figure 20 illustrates the education levels for Nogales, Arizona. All neighborhoods within the city have education levels falling below the 20% mark. Although not illustrated in the map, the average education level for neighborhoods displayed as falling within the first quintile is 11%, while the average for those in the second quintile is 31%. The entire area is well below the mean for the sample, which was 45%.

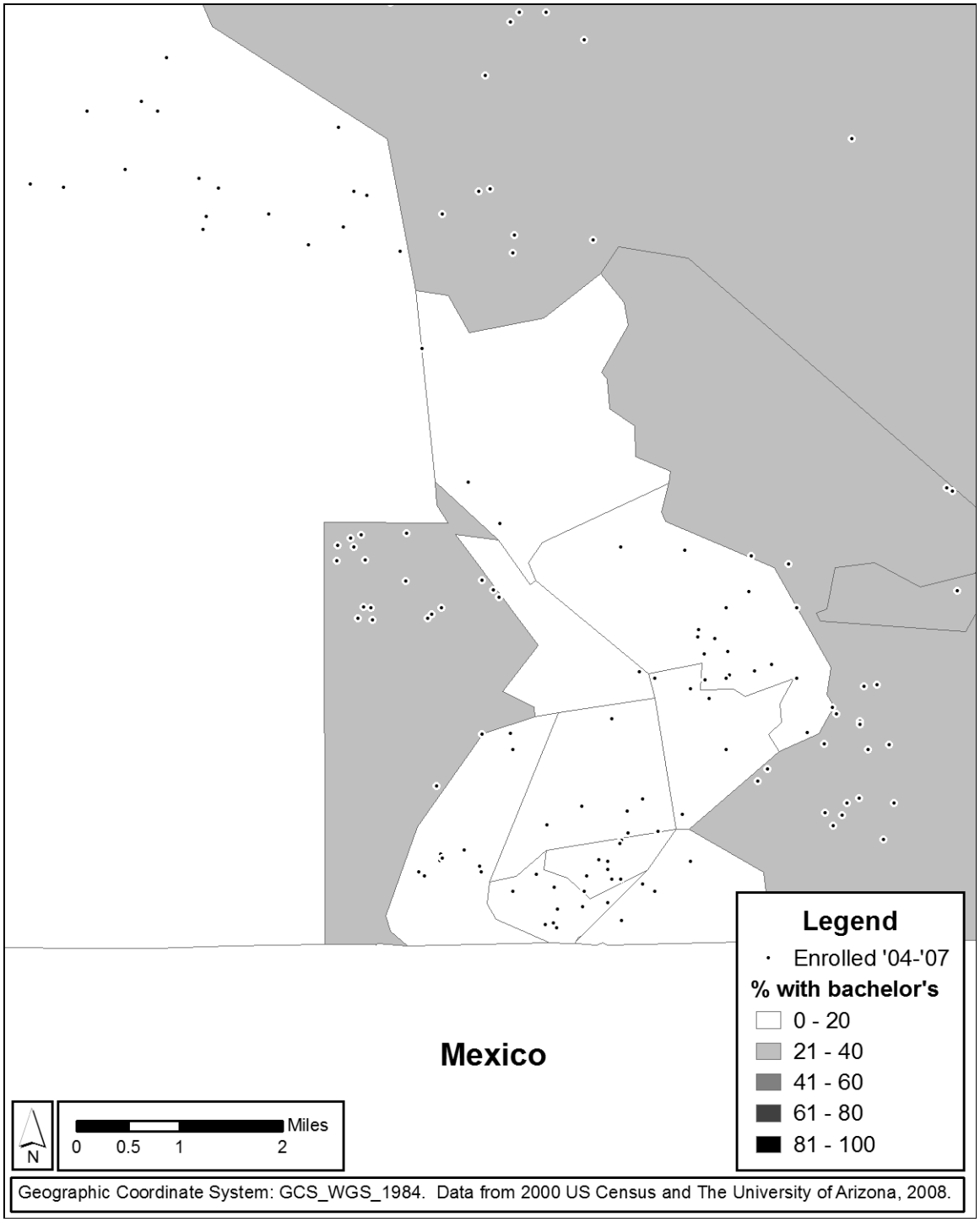

Figure 20. Enrollments with education levels as a percent (bachelor's degree or higher) for block groups for Nogales, Arizona.

Spatial distributions of neighborhood education levels near Los Angeles

Figure 21 serves to illustrate once again the differences between resident students and nonresidents. Enrollees from this area are coming from neighborhoods with the highest education levels, with most coming from neighborhoods in the 61-80% range. As with all maps, there are exceptions, but those from neighborhoods with low education levels are an obvious minority for nonresident students enrolling at the university.

Figure 21. Enrollments with education levels as a percent (bachelor's degree or higher) for block groups for western extent of Los Angeles, California.

Spatial patterns of enrollments by neighborhood racial/ethnic compositions.

I now map the racial/ethnic compositions for the selected areas. Rather than map individual ethnic minority populations, I use an aggregate of the four ethnic minority populations in this study. For this sample the percent of a neighborhood that is composed of ethnic minorities returned a mean of .2437 (SD=.00146). In Figures 22-26, neighborhood ethnic minority percentages from the census are disaggregated in quintiles. Using natural break at each 20% increment in the range contained in the census, the mid-level is 41-60%.

Spatial distributions of neighborhood racial/ethnic compositions for Tucson

In examining Figure 22, it is easily observable that the northern portion of Tucson lacks any concentration of ethnic minority populations. By contrast, the southern portion of Tucson has numerous neighborhoods that are classified in the 5th quintile.

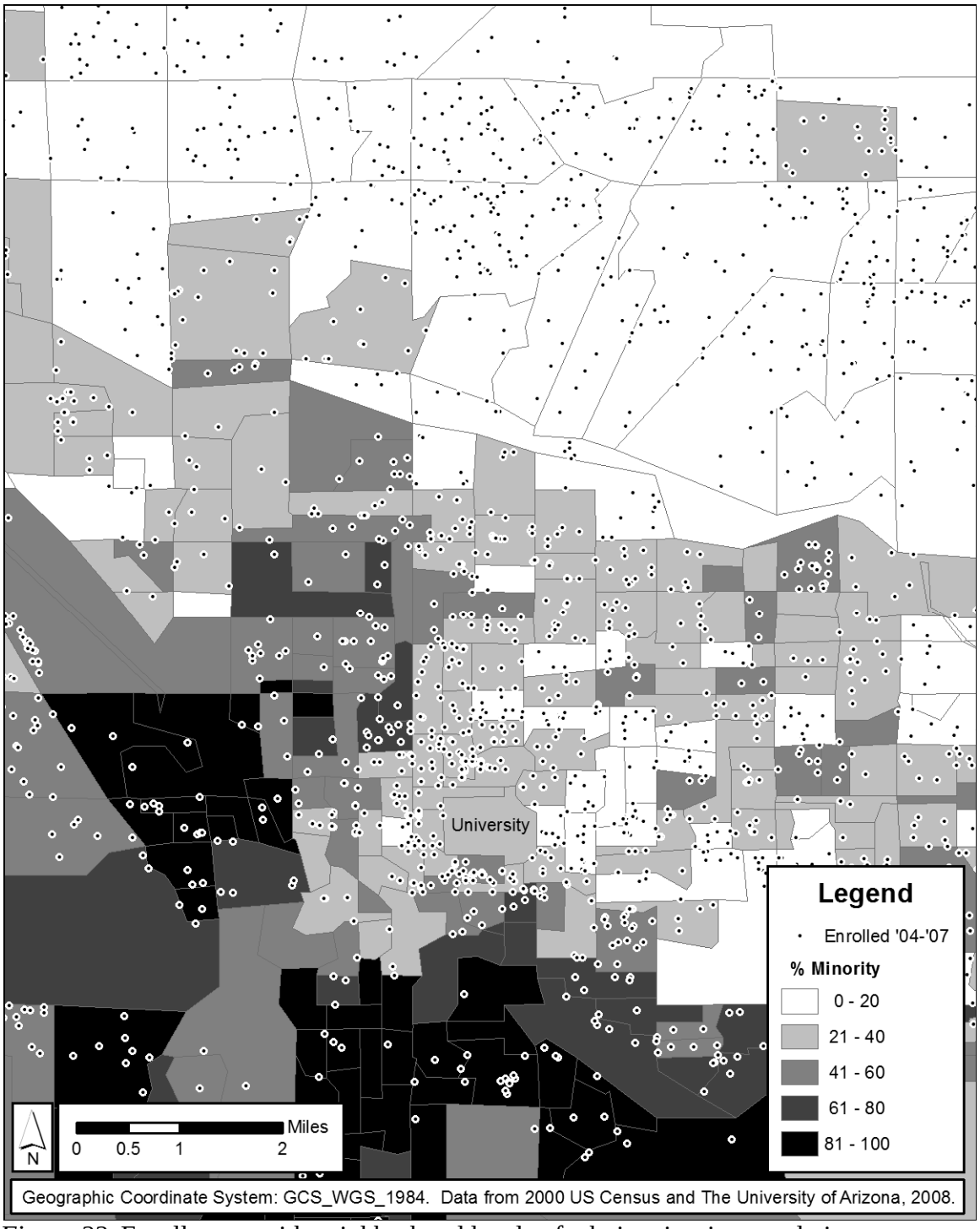

Figure 22. Enrollments with neighborhood levels of ethnic minority population as a percent for block groups for the northern extent of Tucson, Arizona.

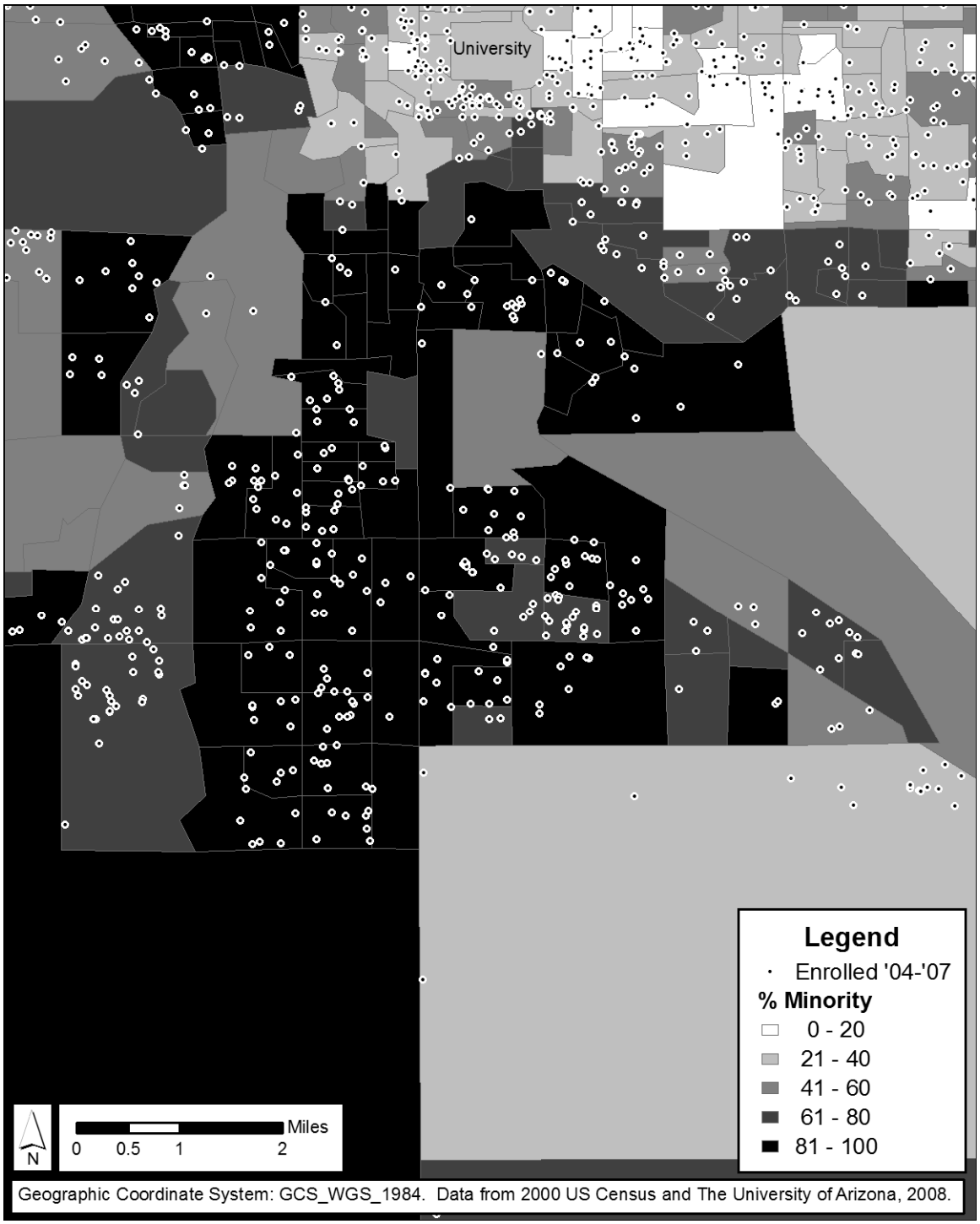

Figure 23. Enrollments with neighborhood levels of ethnic minority population as a percent for block groups for the southern extent of Tucson, Arizona.

Spatial distributions of neighborhood racial/ethnic compositions for Phoenix

In examining Figure 24, it is evident that neighborhoods in the northern area that were previously identified as having high income levels and high education levels, also possesses low levels of ethnic minority populations. By contrast, the southern portion of Phoenix which yields few enrollees for the university has an overwhelmingly high ethnic minority population.

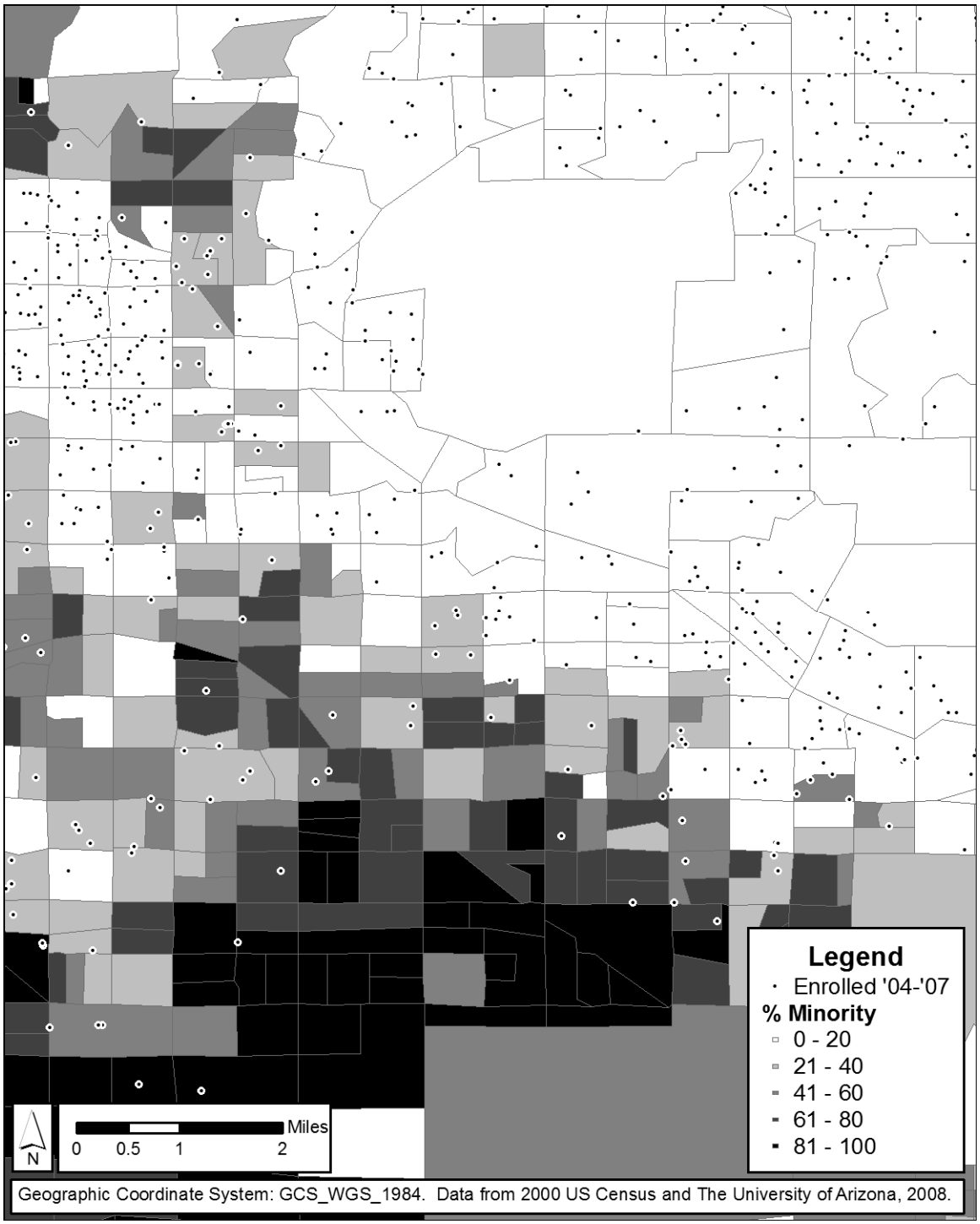

Figure 24. Enrollments with neighborhood levels of ethnic minority population as a percent for block groups for Phoenix, Arizona.

Spatial distributions of neighborhood racial/ethnic compositions for Nogales

In examining Figure 25, the Nogales neighborhoods shown in earlier maps as having low income levels, and low education levels, are identified as having high levels of ethnic minority populations. As with previous maps of Nogales, there is no neighborhood information projected for Mexico.

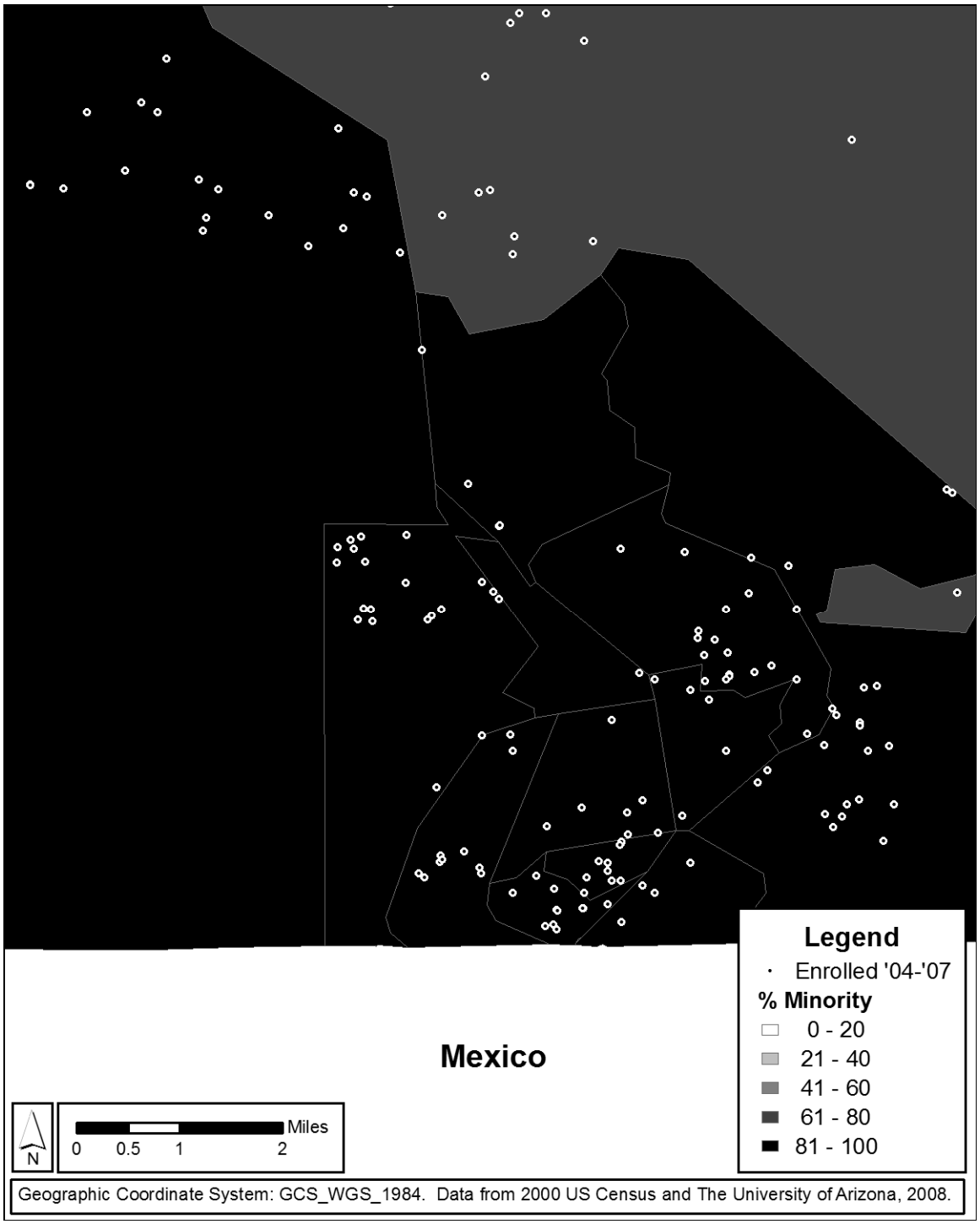

Figure 25. Enrollments with neighborhood levels of ethnic minority population as a percent for block groups for Nogales, Arizona.

Spatial distributions of neighborhood racial/ethnic compositions near Los Angeles

In examining Figure 26, it is clear once again that high income and education levels share a strong relationship a low percentage of ethnic minority population. In this map of the Los Angeles area, the nonresident enrollees are originating from neighborhoods that are predominantly Caucasian. While true, there are also a few enrollees coming from the neighborhoods in the 21-40% range for ethnic minority population.

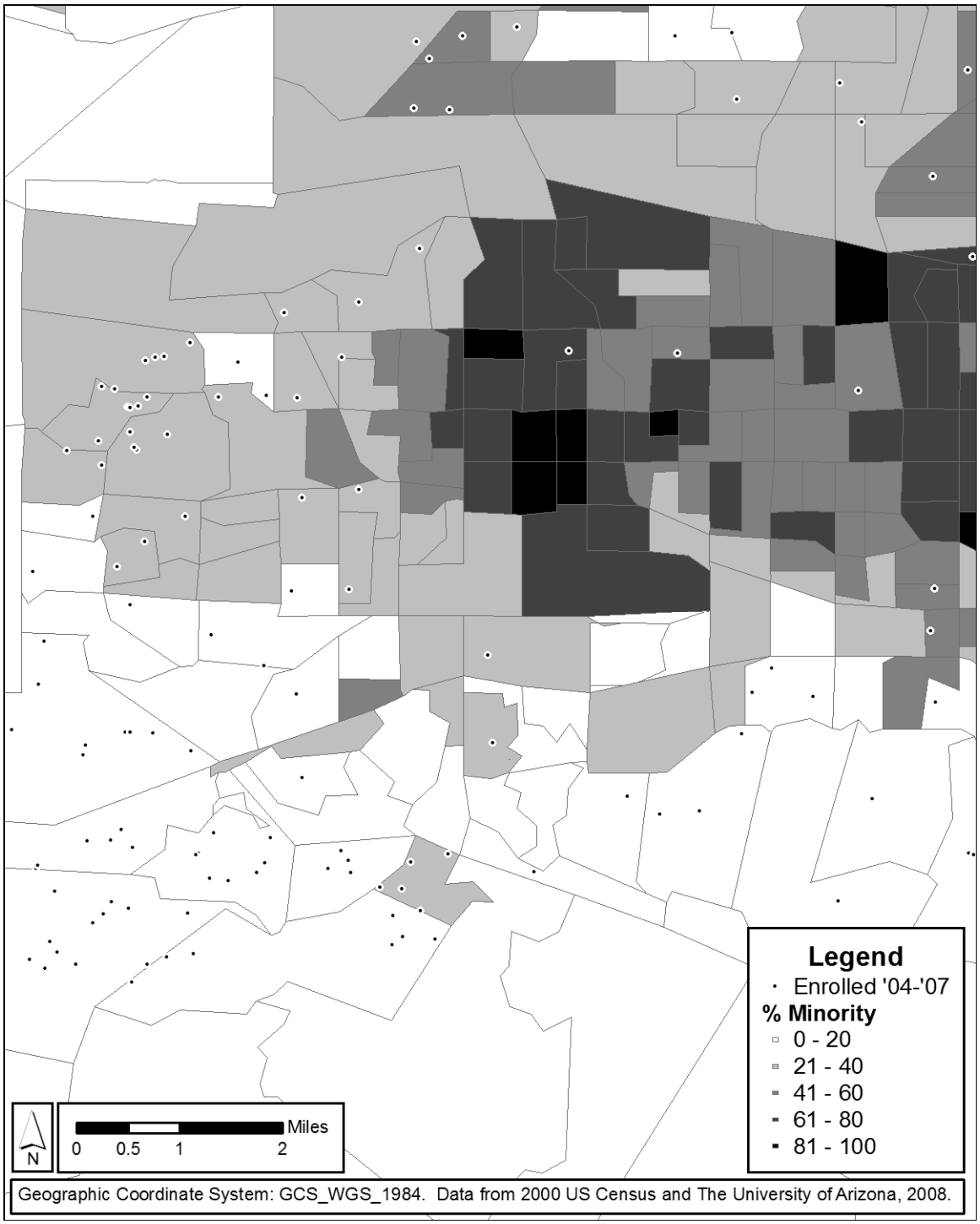

Figure 26. Enrollments with neighborhood levels of ethnic minority population as a percent for block groups for western extent of Los Angeles, California.

Summary of Findings for Research Question 1

In answering the first research question, I presented maps for the entire continental United States, and then focused on four cities that illustrated striking differences in neighborhoods where the university derives its students. These maps also provided a contrast by displaying nearby neighborhoods that did not yield large numbers of enrollees. The maps of Tucson appear to illustrate a diverse community that enrolls students at high levels for the various neighborhood levels of income, education, and ethnic minority population. Conversely, both Phoenix and Los Angeles illustrate differences in enrollment patterns as wealth, education, and the lack of ethnic minority population appear to influence enrollments. Finally, the city of Nogales provided an example of a community with low incomes, low education levels, and a high percent of minorities in the population. Still, Nogales enrolls a substantial number of students at the University of Arizona given its size. These maps do not answer all enrollment questions, but they illustrate how spatial analysis can identify patterns, and facilitate further investigation. This investigation continues in the next section using logistic regression to statistically assess the underlying data that created the patterns evident from my spatial analysis.

Research Question 2: When Using Spatial Analysis, Which Student Characteristics Predict Persistence?

With a visual understanding of the spatial enrollment patterns, I now move to the second research question. Using spatial analysis, what are the characteristics associated with a student's probability to persist? I present multiple models utilizing the 13 student

characteristic variables identified to address this research question. I analyze these variables using a stepwise binomial logistic regression with the dependent variable of year-one retention (RY1). Using the method will enter the variables, and based on their contribution, will ether retain them in the model, or drop them from the model. As mentioned earlier, categorical variables include gender, residency, and race/ethnicity. The equation for my first regression model is:

$$\Pr(RY1 = 1 \mid X_1, X_2, \ldots, X_{13}) = F(\beta_0 + \beta_1(\text{GENDER}) + \beta_2(\text{RACE}) + \beta_3(\text{RESIDENCY}) + \beta_4(\text{GPA}) + \beta_5(\text{PERCENT OF MALES WITH BACHELOR'S}) + \beta_7(\text{PERCENT URBAN AREA}) + \beta_8(\text{PERCENT URBAN CLUSTER}) + \beta_9(\text{COUNT OF TESTS}) + \beta_{10}(\text{APPLICATION DATE IN MONTHS}) + \beta 11(\text{MEDIAN HOUSEHOLD INCOME}) + \beta_{12}(\text{STANDARDIZED TEST SCORE}) + \beta_{13}(\text{EUCLIDIAN DISTANCE FROM THE UNIVERSITY})$$

Equation 1

In Equation 1, Pr is the probability and RY1 is the dependent variable, and X_1-X_{13} are the independent variables. F is the cumulative standard logistic distribution function. In reporting results of the regression, I report findings through the use of two tables. The first table includes those variables retained in the stepwise regression based on statistical significance. The second table in the series reports those variables that were not statistically significant, and as a result were dropped from the regression analysis. The fit of each model is reported using the results of classification tables indicating the accuracy

of my predicted results. The discussion of each result focuses on the odds ratio, an indicator of the odds that an independent variable increases the likelihood that the dependent variable will occur (equal to one). In this analysis, an odds ratio of three indicates that for every unit increase in the independent variable, the odds of year-one persistence increase three times.

In this first regression which includes the entire population (N=21,838), the classification tables returned a result of 80.0 for the percentage correctly predicted. Compared to the starting prediction of 80.1%, this model is close, but does not improve on predicting the outcomes. Findings for RY1 indicate two categorical variables of race/ethnicity were significant and influenced the student's probability to persist. Versus the dummy variable of Caucasian, Asian Americans were 1.339 times more likely to persist while Native Americans were less likely with an odds ratio of .732. The categorical variables of gender and residency were not statistically significant. Six independent variables contributed significantly to the prediction of a student's probability for persisting in their first year of enrollment (Table 24). Five of the variables were not significant, and were dropped from the model (Table 25). In this model, GPA is the strongest predictor with a coefficient of 2.53. Test score is significant, but barely predictive at 1.001. As a proxy for social capital, the percent of males with a bachelor's degree or higher was the second strongest predictor of persistence with a coefficient of 2.230. Percent of urban area also was strong at 1.161. Two cultural capital proxies, number of tests (1.165) and application in months (1.085) also were statistically significant. The economic capital variable of median household income was not

statistically significant. The spatially based distance from the university was not significant.

Table 24
Predictors of RY1 for all enrollees (N=21,838)

Variable	B	S.E.	Sig.	Exp(B)
Race/Ethnicity	—	—	.000	—
Hispanic	.052	0.05	.300	1.053
Asian American	.292	.078	.000	1.339
African American	.126	.091	.168	1.134
Native American	-.312	.097	.001	.732
GPA	.930	0.04	.000	2.534
Percent of bachelor's	.802	.001	.000	2.230
Percent of urban area	.149	.001	.004	1.161
Number of tests	.153	.019	.000	1.165
Application in months	.082	.008	.000	1.085
Standardized Test	.001	0	.000	1.001
Constant	-3.86	0.172	.000	0.021

Table 25
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Residence	.041	1	.839
Gender	.220	1	.639
Percent of urban cluster	.169	1	.681
Median household income	2.288	1	.130
Distance from university	1.186	1	.276

For the regression model using residents, the classification tables returned a result of 80.8% for the percentage correctly predicted. The starting prediction for residents was the same at 80.8%. This model does not improve prediction, but is no worse. Findings for RY1 indicate the categorical variables of race influenced the student's probability to persist. Versus the dummy variable of Caucasian, Asian Americans were 1.562 times more likely to persist while Native Americans were less likely with a beta of .768. The categorical variable of gender was not statistically significant. Six independent variables contributed significantly to the prediction of a student's probability for persisting in their first year of enrollment (Table 26). Five of the variables were not significant, and were therefore dropped from the model (Table 27). In this model, GPA was the second strongest predictor with a coefficient of 2.720, slightly more influential for residents when compared to the entire sample. Test score is significant, but barely predictive at 1.001. The proxy for social capital, the percent of males with a bachelor's degree or higher was the strongest predictor of persistence with a coefficient of 3.702. Two cultural capital proxies, number of tests (1.160) and application in months (1.098) also

were statistically significant. The economic capital variable of median household income was retained in this model, although its influence was not significant. The percent of urban area was dropped from this model, as was gender, percent of urban cluster, and the spatially based distance from the university.

Table 26
Predictors of RY1 for residents (N=14,306)

Variable	B	S.E.	Sig.	Exp(B)
Race/Ethnicity	—	—	.000	—
Hispanic	.141	.059	.018	1.151
Asian American	.446	.100	.000	1.562
African American	.136	.116	.242	1.146
Native American	-.283	.111	.017	.768
GPA	1.001	.052	.000	2.720
Percent of males with bachelor's	1.309	.170	.000	3.702
Number of tests	.149	.027	.000	1.160
Application in months	.093	.011	.000	1.098
Median household income	.000	.000	.014	1.000
Standardized Test	.001	.000	.000	1.001
Constant	-4.549	.215	.000	.011

Table 27
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Gender	.094	1	.759
Percent of urban area	.361	1	.548
Percent of urban cluster	.065	1	.798
Distance from university	.071	1	.790

For the regression model using only nonresidents, the classification tables returned a result of 78.7% for the percentage correctly predicted. The starting prediction was the same at 78.7%. This model does not improve prediction, but as with residents, is inconclusive as to the fit of the model. Findings for RY1 indicate the categorical variables of race did not influence the student's probability to persist at significant levels, nor did the categorical variable of gender. Five independent variables contributed significantly to the prediction of a student's probability for persisting in their first year of enrollment (Table 28). Six of the variables were not significant, and were therefore dropped from the model (Table 29). In this model, GPA was the strongest predictor with a coefficient of 2.192. Percent of urban area was significant and indicated students from urban areas were nearly one and a half times more likely to persist (1.479). As a proxy for social capital, the percent of males with a bachelor's degree or higher was the strongest predictor of persistence with a coefficient of 1.705. Two cultural capital proxies, number of tests (1.161) and application in months (1.060) also were statistically significant. The economic capital variable of median household income was not retained

in this model. The standardized test variable was dropped from this model, as were gender, percent of urban cluster, and the spatially based distance from the university.

Table 28
Predictors of RY1 for nonresidents (N=7,532)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.785	.064	.000	2.192
Percent of Males with Bachelor's	.533	.139	.000	1.705
Percent of urban area	.391	.097	.000	1.479
Number of tests	.150	.029	.000	1.161
Application in months	.058	.014	.000	1.060
Constant	-2.513	.248	.000	.081

Table 29
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Race/Ethnicity	3.972	4	.410
Hispanic	1.062	1	.303
Asian American	.000	1	1
African American	.138	1	.710
Native American	2.674	1	.102
Gender	.345	1	.557
Percent of urban cluster	.040	1	.841
Median household income	.051	1	.821
Standardized test score	.493	1	.483
Distance from university	.836	1	.361

Predictive student characteristics disaggregated by race/ethnicity

With 71% of enrollees classified as Caucasian, it is important to disaggregate by race/ethnicity to isolate the effect of student characteristics on different populations. When using spatial analysis, how do neighborhood racial/ethnic compositions predict the probability that a student will persist? To answer this question, I disaggregate by race and then use the percent of the neighborhood that is comprised by the race/ethnicity disaggregation being analyzed. As an example, using a Hispanic student, I calculate the percent of the individual Hispanic student's block group that is comprised of Hispanic residents. I then use this percentage as a student characteristic variable associated with the individual. Here, I also disaggregate by residency to test if the neighborhood effects of race/ethnicity differ for resident students and nonresident students.

Through the use of associated independent variables, this analysis also will answer two subparts to the second research question: How do neighborhood household income levels predict persistence; and how do neighborhood education levels predict persistence? The equation used is similar to the first equation with slight modifications. Race/ethnicity and residency are removed from the equation as they are used as the filters to select the sample. Also, neighborhood race/ethnic composition is included as an independent variable. In all of the following regressions, the variable [PERCENT RACE/ETHNICITY] is specific to the race/ethnicity being analyzed. The resulting equation for the regression is:

$Pr(RY1 = 1 \mid X_1, X_2,...,X_{13}) = F(\beta_0 + \beta_1(\text{GENDER}) + \beta_2 (\text{PERCENT RACE/ETHNICITY}) + \beta_3 (\text{GPA}) + \beta_4 (\text{PERCENT OF MALES WITH BACHELOR'S}) + \beta_5 (\text{PERCENT URBAN AREA}) + \beta_6 (\text{PERCENT URBAN CLUSTER}) + \beta_7 (\text{COUNT OF TESTS}) + \beta_8 (\text{APPLICATION DATE IN MONTHS}) + \beta 9 (\text{MEDIAN HOUSEHOLD INCOME}) + \beta_{10} (\text{STANDARDIZED TEST SCORE}) + \beta_{11} (\text{EUCLIDIAN DISTANCE FROM THE UNIVERSITY})$

Equation 2

Caucasian student characteristics.

Caucasian students also make up the largest population of resident students at The University of Arizona at 66% (N=9409). From this regression analysis using residents, the classification tables returned a result of 81.7% for the percentage correctly predicted. This is .001 less than the original prediction of 81.8%, indicating the model may not be a good fit. Five student characteristics were significant at the .000 level. Percent of bachelor's was the most influential with a coefficient of 3.250 (Table 30). GPA again proved to be a strong predictor of year-one retention with a coefficient 2.680. The remaining three variables included number of tests (1.140), application in months (1.103) and standardized test score (1.001). Six student characteristics were dropped from the model due to a lack of significance (Table 31).

Table 30
Predictors of RY1 for resident Caucasian students (N=9409)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.986	.066	.000	2.680
Percent of bachelor's	1.179	.138	.000	3.250
Number of tests	.131	.034	.000	1.140
Application in months	.098	.013	.000	1.103
Standardized test score	.001	.000	.000	1.001
Constant	-4.850	.267	.000	.008

Table 31
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Gender	.260	1	.610
Percent of urban area	.364	1	.547
Percent of urban cluster	.145	1	.704
Median household income	3.246	1	.072
Percent Caucasian (neighborhood)	.142	1	.707
Distance from university	.030	1	.863

From this regression analysis of the nonresident Caucasian students (N=6,084), the classification tables returned a result of 79.4% for the percentage correctly predicted, equal to the original prediction of 79.4%, providing inconclusive results to test the fit of this model. Six variables were found significant and were retained in this model (Tables 32-33). The main difference between the resident and nonresident populations appears to be the significance of the percent of a neighborhood that is Caucasian. For nonresidents,

the percent of a neighborhood that is Caucasian has a negative effect on the student's probability to persist. Both GPA (2.208) and Percent of Bachelor's (2.101) remain the strongest predictors of persistence. Percent of urban area has the next the strongest influence with a coefficient of 1.338. The other cultural and social capital proxies, application in months (1.090) and number of tests (1.169) are significant.

Table 32
Predictors of RY1 for nonresident Caucasian students (N=6,084)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.792	.072	.000	2.208
Percent of bachelor's	.742	.175	.000	2.101
Number of tests	.156	.033	.000	1.169
Application in months	.087	.016	.000	1.090
Percent of urban area	.291	.112	.000	1.338
Percent Caucasian (neighborhood)	-.710	.234	.002	.492
Constant	-2.206	.333	.000	.110

Table 33
Predictors of RY1, variables not in the equation for nonresident Caucasian students

Variable	Score	df	Sig.
Gender	.418	1	.518
Percent of urban cluster	.291	1	.589
Standardized test score	1.236	1	.266
Median household income	.232	1	.630
Distance from university	.003	1	.954

Hispanic student characteristics.

Hispanic students make up the largest ethnic minority population at The University of Arizona with 15.78% (N=3,447) in this sample. Of these, 81.2% (N=2798) are resident students. From the regression analysis using Hispanic residents, the classification tables returned a result of 77.4% for the percentage correctly predicted. Without the model, 77.5% of the selected cases were correctly predicted indicating the model may not be a good fit.

More than any group in this study, resident Hispanic students evidenced major differences in odds ratios between resident and nonresident populations. Resident Hispanic students have higher odds of persisting when they come from a neighborhood that has a higher percentage of Hispanic residents. For residents, the influence is stronger for percent Hispanic (neighborhood) with a coefficient of 3.228, than for either GPA (3.047) or Percent of bachelor's (2.790) (Tables 34-35). Also significant for residents were the number of tests (1.159), application in months (1.052) and the standardized test score (1.001).

Table 34
Predictors of RY1 for resident Hispanic students (N=2,798)

Variable	B	S.E.	Sig.	Exp(B)
GPA	1.114	.109	.000	3.047
Percent of bachelor's	1.026	.264	.000	2.790
Number of tests	.147	.063	.020	1.159
Application in months	.050	.023	.027	1.052
Standardized test score	.001	.000	.004	1.001
Percent Hispanic (neighborhood)	1.172	.337	.001	3.228
Constant	-4.525	.453	99.571	.011

Table 35
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Gender	1.533	1	.216
Percent of urban area	.273	1	.601
Percent of urban cluster	.431	1	.511
Median household income	3.671	1	.055
Distance from university	.029	1	.523

From the regression analysis of the nonresident Hispanics in this study (N=649), the classification tables returned a result of 77.2% for the percentage correctly predicted, a substantial improvement over the original prediction of 76.1%. Findings indicate two variables are significant (Tables 36-37). GPA has a coefficient of 3.210 and Percent of bachelor's is 2.969. As opposed to the resident analysis, Percent Hispanic

(neighborhood) does not have a significant contribution to the odds of year-one persistence.

Table 36
Predictors of RY1 for nonresident Hispanic students (N=649)

Variable	B	S.E.	Sig.	Exp(B)
GPA	1.166	.212	.000	3.210
Percent of bachelor's	1.088	.415	.009	2.969
Constant	-2.928	.712	.000	.053

Table 37
Predictors of RY1, variables not in the equation for nonresident Hispanic students

Variable	Score	df	Sig.
Gender	.117	1	.862
Percent of urban area	.308	1	.579
Percent of urban cluster	.039	1	.844
Number of tests	2.564	1	.109
Application in months	.663	1	.415
Standardized test score	.260	1	.610
Median household income	.642	1	.423
Percent Hispanic (neighborhood)	.030	1	.862
Distance from university	.006	1	.733

Asian American student characteristics.

Asian American students made up the second largest ethnic minority population for this sample, accounting for 7.10% of all students (N=1,551). Of these, 73.7% (N=1,143) are resident students. Are they influenced by the same persistence indicators as the majority population? From this regression analysis selecting only residents, the classification tables returned a result of 88.5% for the percentage correctly predicted. This is the same as the beginning prediction for this group, indicating the results are inconclusive as to fit.

Findings for resident African American students at The University of Arizona indicate GPA has the strongest influence (2.629) (Tables 38-39). Standardized test score also influences with a coefficient of 1.002. The influence of the timing of the application, a proxy for cultural and social capital, had a coefficient of 1.193.

Table 38
Predictors of RY1 for resident Asian American students (N=1143)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.967	.199	.000	2.629
Application in months	.177	.048	.000	1.193
Standardized test score	.002	.001	.000	1.002
Constant	-5.300	.851	.000	.005

Table 39
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Gender	8.068	1	.005
Percent of bachelor's	1.185	1	.276
Percent of urban area	.2.263	1	.132
Percent of urban cluster	.915	1	.339
Median household income	.482	1	.488
Percent Asian American. (neighborhood)	.013	1	.908
Distance from university	1.500	1	.221

No variables were found significant for this model using nonresident Asian American students; therefore no results are reported beyond the following two tables (Tables 40-41).

Table 40
Predictors of RY1 for nonresident Asian American students (N=408)

Variable	B	S.E.	Sig.	Exp(B)
Constant	1.350	.122	.000	3.857

Table 41
Predictors of RY1, variables not in the equation for nonresident Asian American students

Variable	Score	df	Sig.
Gender	.050	1	.823
GPA	2.000	1	.157
Percent of bachelor's	.2.042	1	.153
Percent of urban area	.106	1	.745
Percent of urban cluster	.078	1	.780
Number of tests	.007	1	.932
Application in months	.953	1	.329
Standardized test score	1.588	1	.208
Median household income	2.500	1	.114
Percent Asian Americans (neighborhood)	.080	1	.777
Distance from university	.921	1	.337

African American student characteristics.

I begin with African American students. Relative to the entire student population, African American students available for this study comprised 3.49% (N=763) of the enrollment at The University of Arizona. Of these students, 63.8% (N=487) are resident students. Are African American students at The University of Arizona influenced by the same persistence indicators as the majority population? Without the model, the prediction was correct for 76% of the sample. Using this regression model, the classification tables indicate an improvement in the prediction, returning a result of 77.2% for the percentage correctly predicted. Findings for resident African American students at The University of Arizona indicate GPA has the strongest influence (2.939).

Standardized test score also influences with a coefficient of 1.002. Although seemingly small, the difference between the lowest possible (400) test score and the highest possible (1600), would increase a student's odds of persisting to the second year by a factor of 2.4 (Tables 42-43).

Table 42
Predictors of RY1 for resident African American students (N=487)

Variable	B	S.E.	Sig.	Exp(B)
GPA	1.078	.259	.000	2.939
Percent of urban cluster	-1.098	.413	.008	.334
Standardized test score	.002	.001	.006	1.002
Constant	-4.280	.971	.000	.014

Table 43
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Gender	.066	1	.797
Percent of bachelor's	.524	1	.469
Percent of urban area	.114	1	.735
Application in months	.123	1	.726
Percent African American. (neighborhood)	.061	1	.804
Number of tests	.335	1	.563
Median household income	.314	1	.576
Distance from university	.045	1	.831

From this next regression analysis of nonresident African American students (N=276), the classification tables returned a result of 72.8% for the percentage correctly predicted, an improvement over the original prediction of 72.5%. Only two variables were retained in this model (Tables 44-45). The results stand in contrast to the earlier results for the entire population, as neither GPA, nor Percent of Bachelor's is significant. Instead, percent of urban area had the strongest influence at 3.812, and application in months exerted a negative influence with a coefficient of .876.

Table 44
Predictors of RY1 for nonresident African American students (N=276)

Variable	B	S.E.	Sig.	Exp(B)
Percent of urban area	1.338	.5540	.013	3.812
Application in months	-.132	.061	.031	.876
Constant	.614	.609	.314	1.847

Table 45
Predictors of RY1, variables not in the equation for nonresident African American students

Variable	Score	df	Sig.
Gender	.117	1	.862
GPA	.627	1	.429
Percent of bachelor's	.300	1	.584
Percent of urban cluster	.374	1	.541
Number of tests	.675	1	.411
Standardized test score	3.11	1	.577
Median household income	.001	1	.978
Percent African American (neighborhood)	3.110	1	.078
Distance from university	2.535	1	.111

Native American student characteristics.

Native American students make up the smallest ethnic minority with 2.67% of the students in this sample (N=584). Of these Native American students, 80.3% (N=469) are considered resident students. From this regression analysis using residents, the classification tables returned a result of 68.4% for the percentage correctly predicted, an improvement over the original prediction of 67.6% indicating the model is a good fit. GPA is the strongest predictor of persistence. The cultural capital variables of number of tests (1.525) and application in months (1.118) improve the odds of persistence (Tables 46-47). Median household income was significant, although its influence was negligible (1.000). Standardized test scores evidenced a minor influence (1.001). Distance from the university had a slightly negative relationship on the probability to persist with a

coefficient of .998. This may the result of the physical location of two major Native American reservations in the far northeast corner of Arizona, greater than 300 miles from the university campus.

Table 46
Predictors of RY1 for resident Native American students (N=469)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.738	.234	.002	2.091
Number of tests	.422	.126	.001	1.525
Application in months	.112	.049	.021	1.118
Median household income	.000	.000	.025	1.000
Standardized test score	.001	.000	.006	1.001
Distance from university	.-.002	.001	.009	.998
Constant	-2.954	.825	.000	.052

Table 47
Predictors of RY1, variables not in the equation

Variable	Score	df	Sig.
Gender	.001	1	.982
Percent of bachelor's	3.275	1	.070
Percent of urban area	.000	1	.992
Percent of urban cluster	.444	1	.505
Median household income	2.078	1	.149
Standardized test score	.019	1	.889
Percent Native American (neighborhood)	.012	1	.911

From the regression analysis of the nonresident Native Americans in this study (N=115), the classification tables returned a result of 67.0% for the percentage correctly predicted, less accurate than the original prediction of 67.8% indicating the model may not be a good fit. Findings indicate two variables are significant (Tables 48-49). This regression returns the largest influence of GPA compared to all other regressions. A one unit change in GPA makes the student 9.345 times more likely to persist to the second year of enrollment. Also, being from an out-of-state neighborhood that is predominantly Native American has a negative impact upon the likelihood the student will persist. The cultural and social capital variable of number of tests also strongly influenced the odds with a coefficient of 1.880.

Table 48
Predictors of RY1 for nonresident Native American students (N=115)

Variable	B	S.E.	Sig.	Exp(B)
GPA	2.235	.617	.000	9.345
Number of tests	.632	.290	.030	1.880
Percent Native American (neighborhood)	-2.606	.988	.008	.074
Constant	-6.896	1.982	.001	.001

Table 49
Predictors of RY1, variables not in the equation for nonresident Native American students

Variable	Score	df	Sig.
Gender	1.762	1	.184
Percent of bachelor's	1.932	1	.165
Percent of urban area	.017	1	.896
Percent of urban cluster	.232	1	.630
Number of tests	1.739	1	.187
Application in months	.002	1	.964
Standardized test score	.286	1	.593
Median household income	2.188	1	.139
Distance from university	1.840	1	.175

Mapping Probabilities for RY1

Mapping areas most likely to yield persisting Hispanic students (Year one).

With the identification of student characteristics predicting persistence through the regressions, it now becomes possible to spatially identify neighborhoods containing these characteristics. In doing so, recruitment efforts could be targeted by geography, and mailing lists adjusted according to probabilities. As one example, I select Hispanic recruitment as an illustration of the implications for practice. To map these areas for Hispanic students, I create a score using inputs that are predictive. For resident Hispanic students, I use the following to create a score variable:

((P007010/P00701)*3.228)) +(([EDU_Percent] /100)*2.790)

Where P007010 is the total population of Hispanics in the block group, and P007001 is the total population in the block group. This is then multiplied by the result of the logistic regression coefficient associated with an urban area (3.228). As discussed earlier, EDU_Percent is the percent of males in a neighborhood with a bachelor's degree or higher, and 2.790 is the coefficient for Hispanic residents. This creates a variable I term "HispanicResidentScore". This formula alone would not be used in recruitment; instead it would be used with other factors identified in the regression to create a composite score indicating a student's probability for persistence. As many institutions already possess capabilities to rate the prospective student's likelihood of persisting, a derivative of this score could be added to existing prediction formulas (Noel-Levitz, 2009).

I focus the first series of analysis within Arizona, presenting a map of the entire state, and then I present the findings in accordance with the maps already presented for Tucson, Phoenix, and Nogales. Findings are mapped using quintiles. It is important to note, the quintiles are based on odds ratios for neighborhoods only within Arizona. Category selection is based on Jencks natural breaks. Rather than numerical identifiers, I simplify the labels, using the following five categories for the odds of persistence in year one:

- Likely
- More likely
- Average
- Less likely
- Least likely

Areas most likely to yield resident Hispanic students persisting beyond RY1.

Arizona.

In the first map, the entire state of Arizona is projected with areas most likely to yield students who would persist during the first year of enrollment (Figure 27). This is a reflection of the influence of neighborhoods with high Hispanic populations and high education levels. High levels in one category may offset low levels in the other category. An area with low levels in both categories, as is the case with the northeastern section of the map, is devoid of shading. This indicates, ceteris paribus, a Hispanic student from this area would have lower odds of persisting in comparison to a student from the much darker southern section of the map near Nogales. At this scale, it is not possible to see how areas within the Phoenix and Tucson metropolitans fare with respect to the scores.

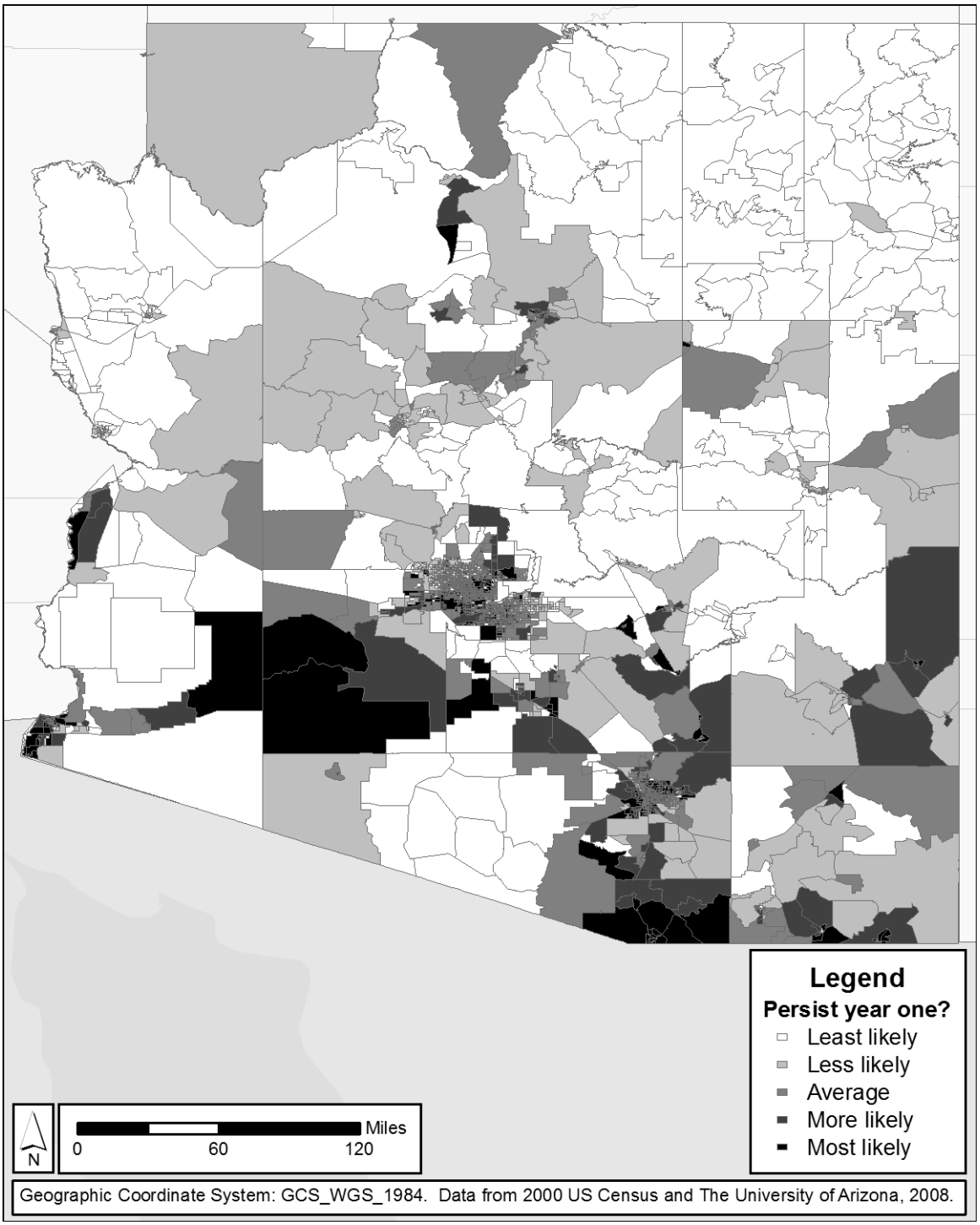

Figure 27. Block groups in Arizona categorized by probability to persist for Hispanic students.

Tucson.

In the second and third maps, the northern and southern extents of Tucson are projected with areas most likely to yield students who would persist during the first year of enrollment (Figures 28-29). Here, it is possible to see how strong influences of educational levels (in the north) or large Hispanic populations can influence the map. Interpreting the results, an enrollment manager might conclude that a student with a high GPA and a high SAT score who had applied in the first month the application was available would have an extremely high probability to persist from nearly all Tucson neighborhoods. As discussed further in Chapter Six, any interpretation would need to be cognizant of the post-enrollment influences, including the potential effects of retention-focused programs at The University of Arizona.

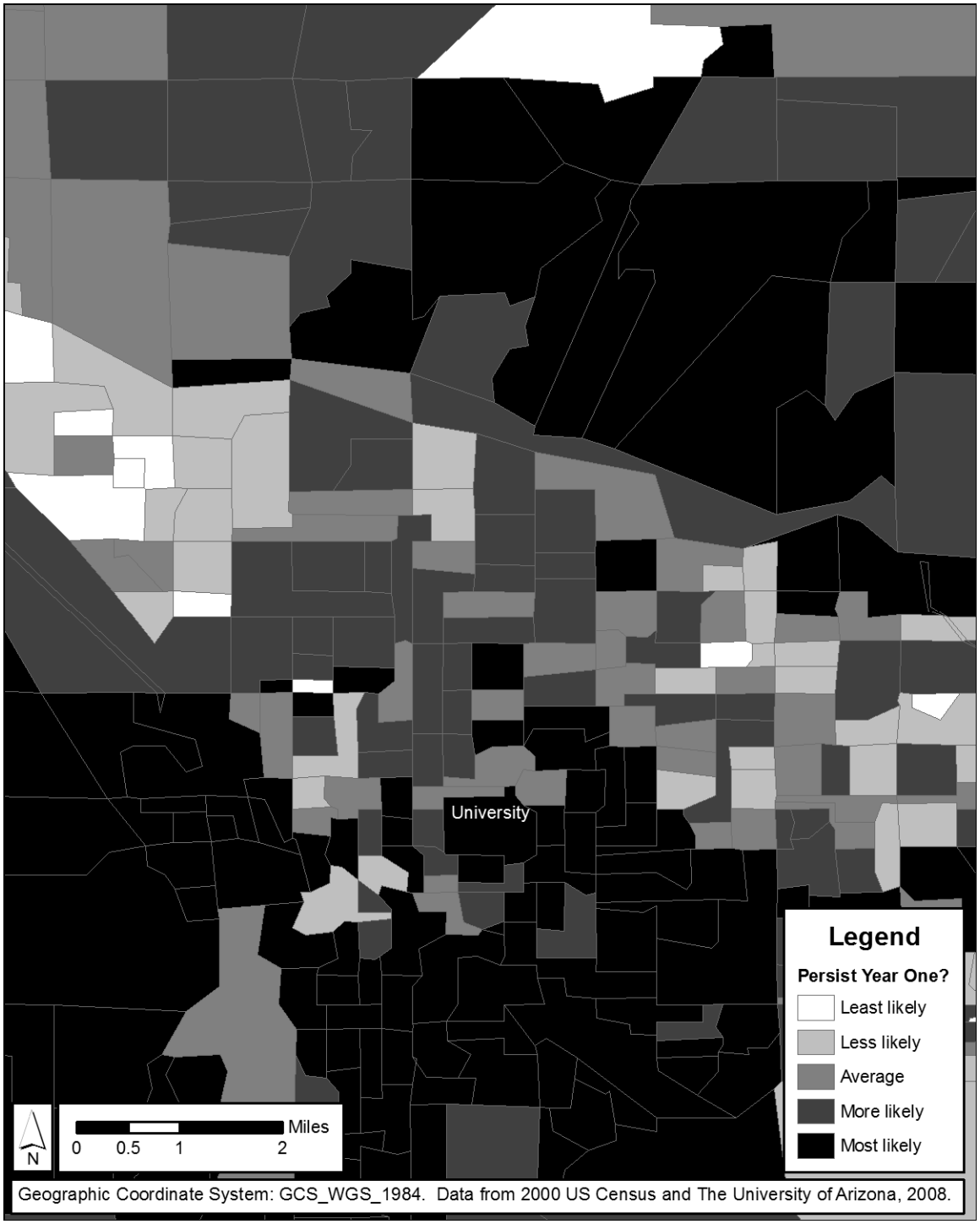

Figure 28. Block groups in Tucson (northern extent) categorized by probability to persist for Hispanic students.

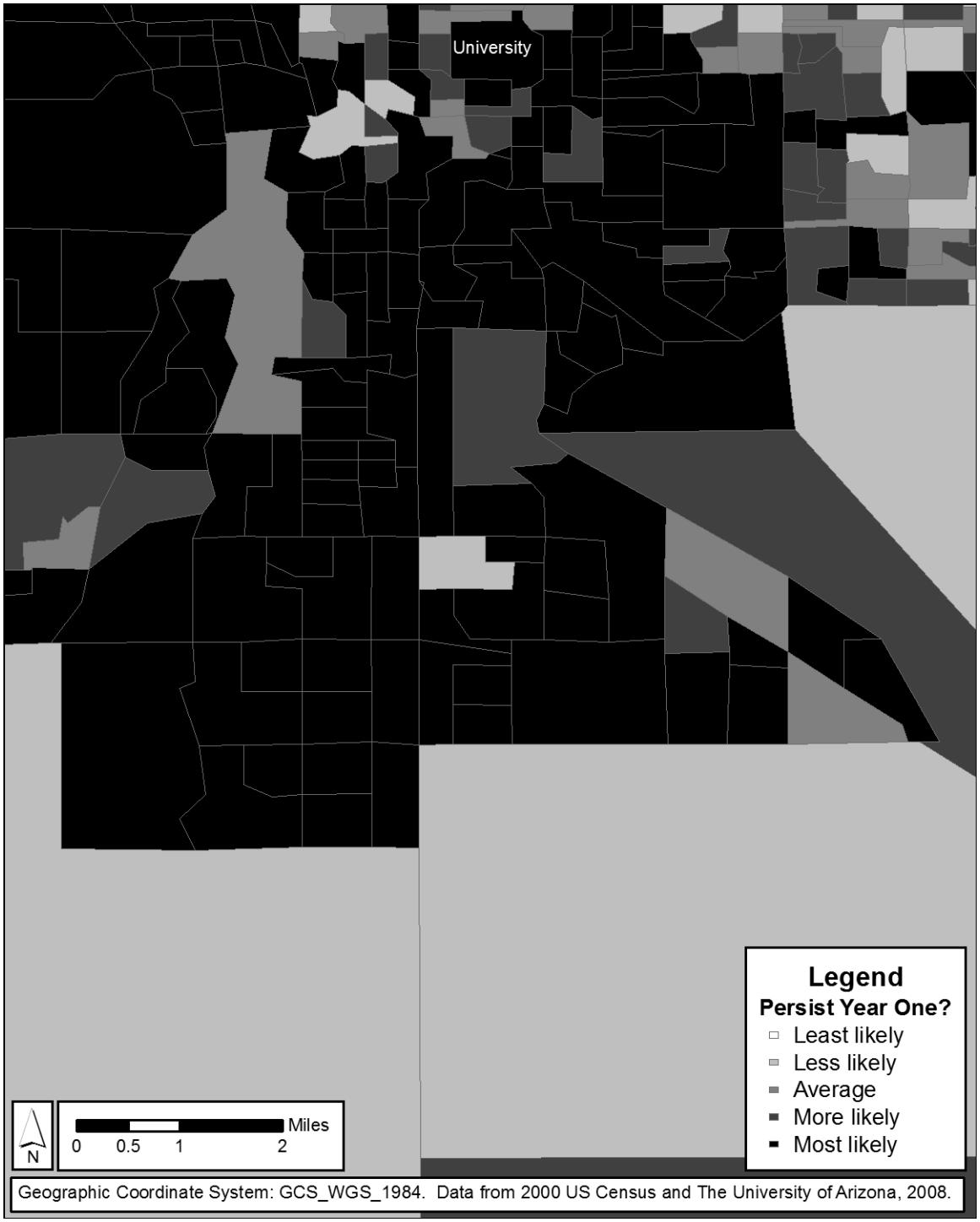

Figure 29. Block groups in Tucson (southern extent) categorized by probability to persist for Hispanic students.

Phoenix.

This map of Phoenix assists in understanding the value of this score (Figure 30). Although the southern area yields few enrollees for the university, the value of the analysis is to provide empirically based evidence that a student from these neighborhoods (with other positive persistence characteristics) would have an excellent chance to persist. At the same time, a student from other areas designated as *Less likely* could be admitted, but could be identified as a student potentially in need of additional institutional resources to persist.

Figure 30. Block groups in Phoenix, Arizona categorized by probability to persist for Hispanic students.

Nogales.

Admittedly, this map provides little detail beyond the blackness indicating all neighborhoods would yield students most likely to persist if they also possessed the academic qualifications necessary (Figure 31). The *Most likely* designation for all neighborhoods on this map is the result of substantially high populations of Hispanic students. However, it serves the purpose of illustrating that if students from these neighborhoods possess adequate academic qualifications, they statistically would be more likely to persist. Recruiting from these areas would make sense for the institution—if the enrollment management goal is to maintain (or increase) institutional persistence rates. Recruiting strong candidates from these neighborhoods also would fulfill enrollment management goals of raising the level of institutional diversity.

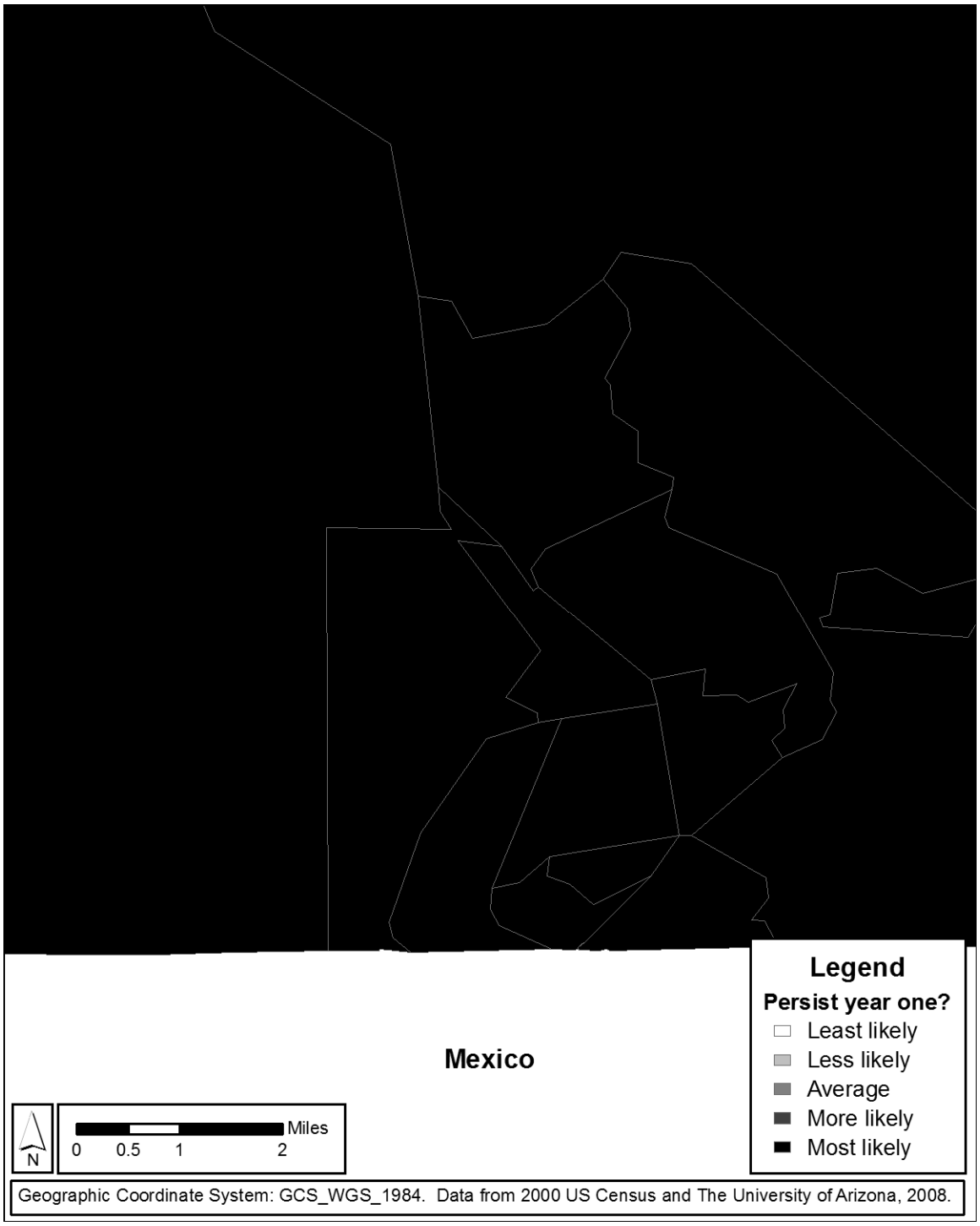

Figure 31. Block groups in Nogales categorized by probability to persist for Hispanic students.

Mapping areas most likely to yield persisting nonresident Hispanic students (Year one).

For nonresident Hispanic students, the odds ratio was created differently, as the percent of the neighborhood comprised of Hispanic residents does not influence the odds of persistence, and the level of education influences nonresidents to a greater degree. For nonresident Hispanics, I use the following formula to create an odds variable

(([EDU_Percent] /100)*2.969)

California.

The map of California serves as a useful tool in identifying areas where the recruitment of Hispanic students statistically would yield the greatest results from a persistence point of view (Figure 32). For Hispanic students, this map is simply a reflection of those neighborhoods in California with the highest levels of education for males. It does illustrate the concentrations of neighborhoods where enrollment managers unfamiliar with the state might want to use the information to tailor mailing lists, contact local high schools, or participate in selected college fairs. It could be equally useful to post-admission efforts to improve persistence by identifying students who may benefit from additional institutional resources.

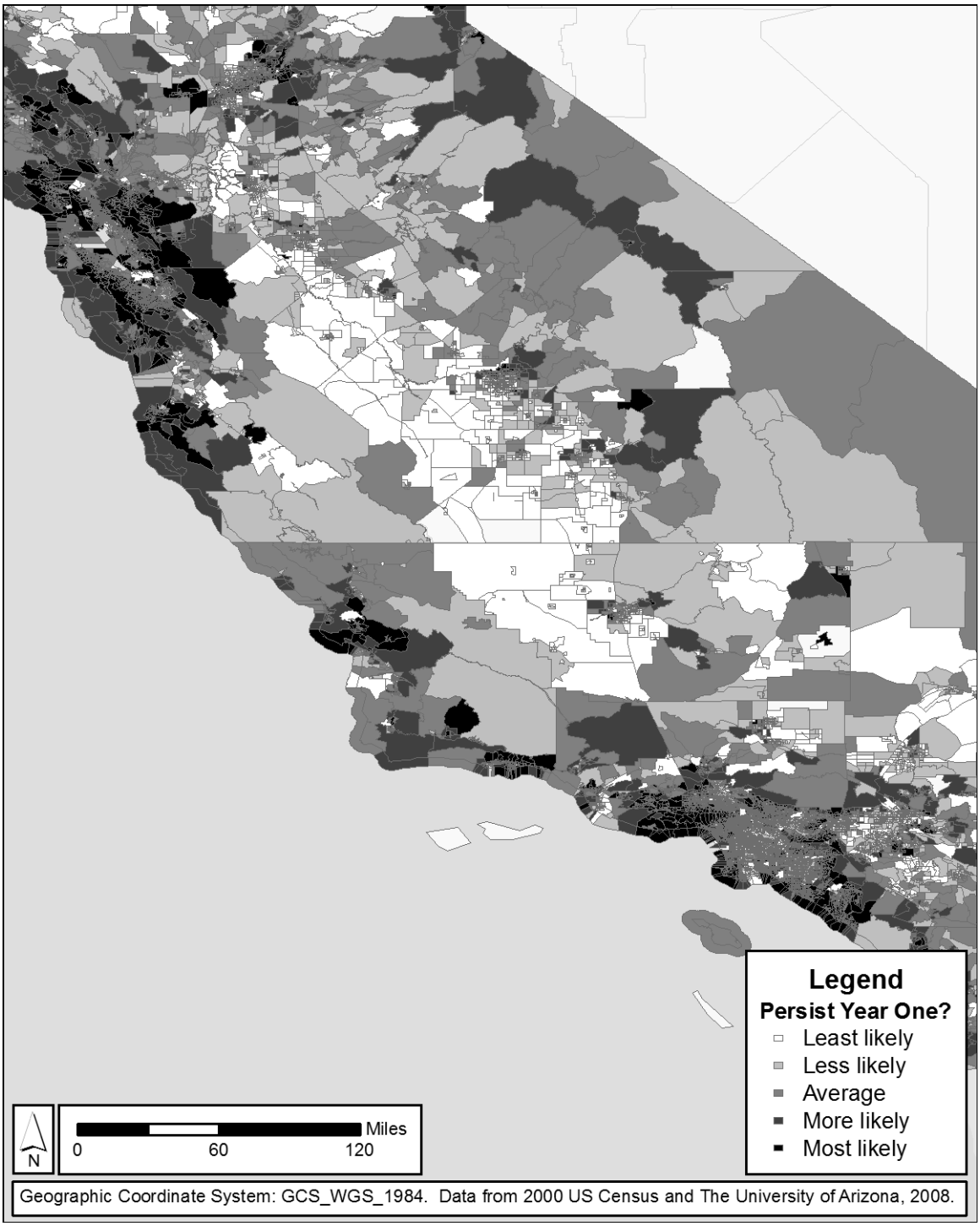

Figure 32. Block groups in California categorized by probability to persist for Hispanic students.

Los Angeles.

This map of neighborhoods west of Los Angeles indicates nonresident students from lower portions of the map would have high probabilities to persist (Figure 33). Students from the inner areas of the maps, identified as *Less likely* and *Average* for persisting. Again, rather than a rationale to decline admission to these students, this information can be used as a tool to provide advanced knowledge to administrators—students admitted from the *Least likely* neighborhoods statistically are in need of additional resources to improve their chances of persisting.

Figure 33. Block groups in Los Angeles categorized by probability to persist for Hispanic students.

Summary of Findings for Research Question 2.

In answering this second research question, I presented the results of regression analysis of 13 student characteristics and their relationship to the dependent variable for year-one persistence. Analysis was done for the entire sample, and then disaggregated by resident and nonresident enrollees. For all three, it was clear that GPA is the most predictive of student persistence during the first year. Also, cultural and social capital proxies showed evidence of their value in predicting persistence, most notably the neighborhood educational level of males. Additionally, the number of tests taken and the timing of the admission application were significant findings. Although significant for the entire sample, and for nonresidents, percent of urban area was not significant for resident students.

From these initial findings, I then disaggregated by race, again analyzing resident and nonresidents as separate populations. In doing so, the analysis of resident populations found GPA to be significant for all race/ethnicity categories. Educational levels were important, but only for the two largest race/ethnicity categories, Caucasian and Hispanic. For Caucasians it was more predictive than GPA. One final cultural and social capital variable, the number of tests taken proved to be significant for Caucasians, Hispanics, and Native Americans. For Hispanics, the percent of their neighborhood that was Hispanic proved to be a significant predictor of persistence. This was the only race/ethnicity where this was the case.

For nonresident students, a smaller population, significant findings were found for all race/ethnicity categories except for Asian Americans. GPA was predictive, but only for Caucasian, Hispanic, and Native American populations. In the case of nonresident Native Americans, the coefficient of GPA was 9.345 and significant at the .000 level. African Americans were most influenced by percent of urban area, and were the only nonresident group where standardized test scores were significant. Now having answered the questions identifying which characteristics are associated with a student's probability to persist, I move on to the final research question.

Research Question 3: When Using Spatial Analysis, Which Student Characteristics Predict Graduation?

With a visual understanding of the spatial enrollment patterns, and the analysis of student characteristics predicting persistence, I address the third research question. Using spatial analysis, what are the characteristics associated with a student's probability to graduate? As I did for the second research question, I use multiple models including the 13 student characteristic variables identified to address this research question. I use a stepwise binomial logistic regression this time with the dependent variable of year-four graduation (or persistence after four years) (RY4G). Categorical variables include gender, residency, and race/ethnicity. The equation for my model is:

$Pr(RY4G = 1 \mid X_1, X_2,...,X_{13}) = F(\beta_0 + \beta_1(GENDER) + \beta_2 (RACE) + \beta_3 (RESIDENCY) + \beta_4 (GPA) + \beta_5 (PERCENT\ OF\ MALES\ WITH\ BACHELOR'S) + \beta_7 (PERCENT\ URBAN\ AREA) + \beta_8 (PERCENT\ URBAN\ CLUSTER) + \beta_9 (COUNT\ OF\ TESTS) + \beta_{10} (APPLICATION\ DATE\ IN\ MONTHS) + \beta 11 (MEDIAN\ HOUSEHOLD\ INCOME) + \beta_{12} (STANDARDIZED\ TEST\ SCORE) + \beta_{13} (EUCLIDIAN\ DISTANCE\ FROM\ THE\ UNIVERSITY))$

Equation 3

Where Pr is the probability and RY4 is the dependent variables, and X_1-X_{13} are the independent variables. F is the cumulative standard logistic distribution function. As with the previous research question, the fit of each model is reported using classification tables indicating the accuracy of predicted results. After comparing to the original prediction without the model, the results are once again evaluated for goodness of fit.

As this regression is examining student characteristics related to graduation or persistence after four years of enrollment, only the enrollees from 2004 are in the sample (N=5,043). In this first regression which includes the entire population, the classification tables returned a result of 67% for the percentage correctly predicted. Compared to the starting prediction of 63.6%, this model substantially improves on predicting the outcomes. Findings for RY4G indicate the categorical variables of race influenced the student's probability to persist. Versus the dummy variable of Caucasian, all populations were slightly less likely to graduate, while Native Americans were only half as likely to

graduate (Tables 50-51). The percent of males with bachelor's degrees or higher improved the odds of graduation by nearly three times, and was the most influential of all the variables, followed by GPA (2.590), number of tests (1.164), application in months (1.113) and standardized tests (1.001).

Table 50
Predictors of RY4G for all 2004 enrollees (N=5,043)

Variable	B	S.E.	Sig.	Exp(B)
Race/Ethnicity	—	—	.007	—
Hispanic	-.155	.090	.085	.857
Asian American	-.126	.124	.309	.882
African American	-.142	.173	.414	.868
Native American	-.713	.205	.001	.490
GPA	.952	.072	.000	2.590
Percent of bachelor's	1.041	.148	.000	2.831
Number of tests	.152	.034	.000	1.164
Application in months	.107	.017	.000	1.113
Standardized test scores	.001	.000	.000	1.001
Distance from university	.000	.000	.018	1.000
Constant	-5.017	.315	.000	.007

Table 51
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Residence	1.182	1	.277
Gender	.008	1	.930
Percent of urban area	2.295	1	.130
Percent of urban cluster	2.029	1	.154
Median household income	.065	1	.799

For the regression model using residents (N=3,464), the classification tables returned a result of 68.7% for the percentage correctly predicted. The starting prediction for residents was four percentage points lower at 64.8%, indicating that this model is a good fit. Since resident students accounted for 70% of the sample, results were similar to the findings for the entire population. Findings for RY4G indicate the categorical variables of race influenced the student's probability to persist. Versus the dummy variable of Caucasian, only Asian American students were close to having the same probability to graduate (.981) (Tables 52-53). As with the first regression for RY4G, Native American students were only half as likely to graduate. The influence of other student characteristics remained in the same order, with similar odds: percent of males with bachelor's or higher (2.629), GPA (2.521), number of tests (1.246), application in months (1.246), and standardized tests (1.001).

Table 52
Predictors of RY4G for residents (N=3,464)

Variable	B	S.E.	Sig.	Exp(B)
Race/Ethnicity	—	—	.026	—
Hispanic	-.131	.101	.193	.877
Asian American	-.019	.151	.900	.981
African American	-.178	.216	.409	.837
Native American	-.714	.224	.001	.490
GPA	.925	.091	.000	2.521
Percent of bachelor's	.967	.186	.000	2.629
Number of tests	.220	.046	.000	1.246
Application in months	.119	.021	.000	1.246
Standardized Test	.001	.000	.000	1.001
Constant	-5.589	.379	.000	.004

Table 53
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Gender	.010	1	.919
Percent of urban area	.940	1	.332
Percent of urban cluster	.356	1	.551
Median household income	1.845	1	.174
Distance from university	.921	1	.337

For the regression model using 1,579 nonresidents, the classification tables returned a result of 63.9% for the percentage correctly predicted. The starting prediction for non residents was the lower at 60.8%. This model improves prediction, and indicates

a good fit. In these findings, race/ethnicity was no longer significant, most likely due to a low number of nonresident ethnic minority students in the sample (Tables 54-55). Four student characteristics were significant at the 95% confidence interval or higher. Of these, percent of males with a bachelor's or higher remained the most influential (2.871), followed by GPA (2.644). Unlike the results for the entire population, and the resident population, nonresident graduates had their odds ratio increased by the percentage of a neighborhood classified as an urban area (1.479). The timing of an application for admission remained important, although it was an event that had occurred four to five years before graduation.

Table 54
Predictors of RY4G for nonresidents (N=1,579)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.972	.118	.000	2.644
Percent of bachelor's	1.055	.252	.000	2.871
Percent of urban area	.391	.097	.000	1.479
Application in months	.090	.027	.001	1.094
Distance from university	.000	.000	.040	1.000
Constant	-3.719	.458	.000	.024

Table 55
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Race/Ethnicity	3.342	4	.502
Hispanic	.106	1	.745
Asian American	2.393	1	.122
African American	.037	1	.847
Native American	.752	1	.386
Gender	.211	1	.646
Percent of urban area	1.412	1	.235
Percent of urban cluster	3.254	1	.071
Number of tests	.166	1	.683
Median household income	.266	1	.606
Standardized Test	.565	1	.452

Predictive student characteristics disaggregated by race/ethnicity

I now disaggregate by race/ethnicity to isolate the effect of student characteristics on the different populations. When using spatial analysis, what are the characteristics associated with a student's probability to graduate, specifically racial/ethnic compositions that predict persistence? As with persistence, in answering this research question, I disaggregate by race and then use the percent of the neighborhood that is comprised by the race/ethnicity disaggregation being analyzed. Also, I again disaggregate by residency to test if the neighborhood effects of race/ethnicity differ for resident students and nonresident students. The equation used is similar to the second equation with modifications to the dependent variable. The resulting equation for the regression is:

$Pr(RY4G = 1 \mid X_1, X_2,...,X_{13}) = F(\beta_0 + \beta_1(GENDER) + \beta_2$ (PERCENT RACE/ETHNICITY) + β_3 (RESIDENCY) + β_4 (GPA)+ β_5 (PERCENT OF MALES WITH BACHELOR'S) + β_7 (PERCENT URBAN AREA) + β_8 (PERCENT URBAN CLUSTER) + β_9 (COUNT OF TESTS) + β_{10} (APPLICATION DATE IN MONTHS)+ β11 (MEDIAN HOUSEHOLD INCOME)+ β_{12} (STANDARDIZED TEST SCORE) + β_{13} (EUCLIDIAN DISTANCE FROM THE UNIVERSITY))

Equation 4

Caucasian student characteristics.

As Caucasian students made up the largest population of first-year resident students at 66% (N=9,409), they also accounted for a similar percentage of graduates (and those still enrolled) after four years at 67% (N=2,323). From the regression analysis using residents, the classification tables returned a result of 69.8% for the percentage correctly predicted. This is 1.5 percentage points better than the original prediction of 68.3%, indicating the model is a good fit. Five student characteristics were significant. Percent of bachelor's was the most influential with a coefficient of 3.088 (Tables 56). GPA again proved to be a strong predictor of graduation with a coefficient 2.653. The remaining three variables included number of tests (1.188), application in months (1.143) and standardized test score (1.001). Six student characteristics were dropped from the model due to a lack of significance (Table 57).

Table 56
Predictors of RY4G for resident Caucasian students (N=2,323)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.976	.113	.000	2.653
Percent of bachelor's	1.127	.233	.000	3.088
Number of tests	.173	.057	.003	1.188
Application in months	.134	.026	.000	1.143
Standardized test score	.001	.000	.000	1.001
Constant	-5.988	.474	.000	.003

Table 57
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Gender	.022	1	.882
Percent of urban area	.324	1	.569
Percent of urban cluster	.155	1	.694
Median household income	1.663	1	.197
Distance from university	.654	1	.419
Percent Caucasian (neighborhood)	.113	1	.737

From this regression analysis of the 1,303 nonresidents, the classification tables returned a result of 64.8% for the percentage correctly predicted, an improvement over the original prediction of 62.2%. Three variables were found significant and were retained in this model (Tables 58-59). Both GPA (2.917) and Percent of Bachelor's remain the strongest predictors of persistence (2.771). The cultural capital proxy, application in months (1.102) remained significant.

Table 58
Predictors of RY4G for nonresident Caucasian students (N=1,303)

Variable	B	S.E.	Sig.	Exp(B)
GPA	1.071	.132	.000	2.917
Percent of bachelor's	1.019	.288	.000	2.771
Application in months	.097	.031	.002	1.102
Constant	-4.203	.521	.000	.015

Table 59
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Gender	.357	1	.550
Percent of urban area	2.008	1	.156
Percent of urban cluster	2.819	1	.093
Number of tests	1.331	1	.249
Median household income	.274	1	.601
Standardized test score	2.044	1	.153
Distance from university	2.677	1	.102
Percent Caucasian (neighborhood)	.248	1	.619

Hispanic student characteristics.

Hispanic students make up the largest ethnic minority population from 2004 enrollees who were still enrolled or had graduated after four years. There were 779 in this sample. Of these, 85.9% (N=669) are resident students. Without the model, the prediction was correct for 56.1% of the sample. From this regression analysis, the results from the classification tables indicate the model improved the prediction, returning a

result of 62.8% for the percentage correctly predicted. Three student characteristics were found to influence prediction. GPA was by far the strongest indicator of graduation (3.044), along with the cultural capital proxies of number of tests (.1.487), and application in months (1.118) (Tables 60-61).

Table 60
Predictors of RY4G for resident Hispanic students (N=669)

Variable	B	S.E.	Sig.	Exp(B)
GPA	1.113	.192	.000	3.044
Number of tests	.397	.108	.000	1.487
Application in months	.112	.048	.019	1.118
Constant	-5.097	.699	.000	.006

Table 61
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Gender	.169	1	.681
Percent of bachelor's	3.655	1	.056
Percent of urban area	1.789	1	.181
Percent of urban cluster	.963	1	.326
Median household income	.787	1	.375
Standardized test score	.907	1	.341
Distance from university	.272	1	.602
Percent Hispanic (neighborhood)	.437	1	.508

From this regression analysis using nonresidents, the classification tables returned a result of 66.4% for the percentage correctly predicted, a substantial improvement over the original prediction of 59.1%. Findings for nonresident Hispanic students indicated that GPA (2.932) was the only student characteristics significantly predicting the student's probability to graduate (Table 62-63).

Table 62
Predictors of RY4G for nonresident Hispanic students (N=110)

Variable	B	S.E.	Sig.	Exp(B)
GPA	.872	.437	.046	2.932
Constant	-2.325	1.357	.087	.098

Table 63
Predictors of RY4G, variables not in the equation for nonresident Hispanic students (N=622)

Variable	Score	df	Sig.
Gender	.409	1	.523
Percent of bachelor's	.694	1	.405
Percent of urban area	.058	1	.809
Percent of urban cluster	2.351	1	.125
Number of tests	.054	1	.817
Application in months	.274	1	.601
Standardized test score	1.097	1	.295
Median household income	.337	1	.561
Percent Hispanic (neighborhood)	.033	1	.856
Distance from university	.054	1	.816

Asian American student characteristics.

Relative to the entire student population of graduates from the 2004 cohort, Asian American students comprise a small portion of all students at The University of Arizona. In the year for which four-year graduation/persistence outcomes were available, only 360 students were available for this analysis. Of these students, 73.9% (N=266) are resident students. Without the model, the prediction was correct for 71.8% of the sample. From this regression analysis, the results from the classification tables indicate the model slightly improved the prediction, returning a result of 73.3% for the percentage correctly predicted. Two student characteristics were found to influence prediction. Standardized test scores were significant at .000 and increased the probability (1.004) of graduating

(Tables 64-65). The percent of the enrollee's neighborhood that was Asian American dramatically decreased the student's probability to graduate.

Table 64:
Predictors of RY4G for resident Asian American students (N=266)

Variable	B	S.E.	Sig.	Exp(B)
Standardized test score	.004	.001	.000	1.004
Percent Asian American (neighborhood)	-7.323	3.049	.016	.001
Constant	-3.138	1.008	.002	.043

Table 65
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Gender	1.813	1	.178
GPA	.023	1	.879
Percent of bachelor's	1.785	1	.182
Percent of urban area	.447	1	.504
Percent of urban cluster	.039	1	.844
Number of tests	.662	1	.416
Application in months	.152	1	.696
Median household income	1.027	1	.488
Distance from university	.024	1	.876

From this regression analysis selecting nonresidents, the classification tables returned a result of 63.8% for the percentage correctly predicted. With a beginning prediction of 54.3% for this group, the results indicate the model has a good fit. Findings

for nonresident Asian American students at The University of Arizona indicate that the timing of the student's application for admission predicted graduation better than other factors (1.313), although the significance was .019 (Tables 66-67). The distance from the university had a slightly negative impact on graduation (.999).

Table 66
Predictors of RY4G for nonresident Asian American students (N=94)

Variable	B	S.E.	Sig.	Exp(B)
Application in months	.272	.116	.019	1.313
Distance from university	.000	.000	.006	.999
Constant	-1.081	.873	.216	.339

Table 67
Predictors of RY4G, variables not in the equation for nonresident Asian American students (N=94)

Variable	Score	df	Sig.
Gender	.032	1	.859
GPA	.818	1	.366
Percent of bachelor's	1.280	1	.258
Percent of urban area	.056	1	.813
Percent of urban cluster	.482	1	.488
Number of tests	.640	1	.424
Standardized test score	1.847	1	.174
Median household income	.400	1	.527
Percent Asian American. (neighborhood)	1.079	1	.299

African American student characteristics.

Relative to the entire student population of graduates from the 2004 cohort, African American students comprise a small portion at The University of Arizona. In the year for which four-year graduation/persistence outcomes were available, only 161 students were available for this analysis. Of these students, 64.6% (N=104) are resident students. Are African American students at The University of Arizona influenced by the same graduation indicators as the majority population? Without the model, the prediction was correct for 51% of the sample. From this regression analysis, the classification tables improved the prediction, returning a result of 57.7% for the percentage correctly predicted. Findings for resident African American students at The University of Arizona identified GPA as the only significant variable that was important in predicting persistence at 2.968 (Tables 68-69). This indicates that a resident African American student entering The University of Arizona with a 3.5 GPA is three times more likely to graduate when compared to an entering resident African American student with a 2.5 GPA.

Due to the small sample size (57), the analysis for nonresident African American graduates did not yield significant results.

Table 68
Predictors of RY4G for resident African American students (N=104)

Variable	B	S.E.	Sig.	Exp(B)
GPA	1.088	.434	.012	2.968
Constant	-3.504	1.428	.014	.030

Table 69
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Gender	.423	1	.515
Percent of bachelor's	.368	1	.544
Percent of urban area	1.615	1	.204
Percent of urban cluster	2.061	1	.151
Number of tests	.084	1	.772
Application in months	.240	1	.624
Median household income	.000	1	.994
Standardized test score	.322	1	.570
Distance from university	.232	1	.630
Percent African American. (neighborhood)	3.316	1	.069

Native American student characteristics.

Native American students make up the smallest ethnic minority population from 2004 enrollees who were still enrolled or had graduated after four years. There were 117 in this sample. Of these, 87.2% (N=102) are resident students. Without the model, the prediction was correct for 60.8% of the sample. From this regression analysis, the results from the classification tables indicate the model improved the prediction, returning a

result of 68.6% for the percentage correctly predicted. Only two student characteristics were found to influence prediction, and both serve as proxies for cultural and social capital. These student characteristics were the number of tests taken by the student (1.659) and the timing of the submission of the application (1.314) (Tables 70-71).

Table 70
Predictors of RY4G for resident Native American students (N=102)

Variable	B	S.E.	Sig.	Exp(B)
Number of tests	.506	.237	.033	1.659
Application in months	.273	.119	.022	1.314
Constant	-3.628	1.106	.001	.027

Table 71
Predictors of RY4G, variables not in the equation

Variable	Score	df	Sig.
Gender	.467	1	.494
GPA	.834	1	.361
Percent of bachelor's	.582	1	.446
Percent of urban area	.786	1	.375
Percent of urban cluster	.206	1	.650
Median household income	.489	1	.485
Standardized test score	.395	1	.530
Percent Native American (neighborhood)	.316	1	.574
Distance from university	.298	1	.585

Table 72
Predictors of RY4G, variables not in the equation for nonresident Native American students

Variable	Score	df	Sig.
Gender	.032	1	.858
GPA	2.014	1	.156
Percent of bachelor's	2.358	1	.125
Percent of urban area	.077	1	.782
Percent of urban cluster	.944	1	.331
Number of tests	.376	1	.540
Application in months	.111	1	.740
Standardized test score	2.643	1	.104
Median household income	.008	1	.927
Percent Native American. (neighborhood)	.328	1	.567
Distance from university	.008	1	.927

Conclusion

This chapter began with a descriptive analysis of The University of Arizona to provide greater context for understanding the findings of this study. I then presented the demographic and geodemographic information relating to the students included in this sample. Next, I presented findings describing in detail the dependent and independent variables used for analysis. Lastly, the findings of this research were presented as they relate to each of the three main research questions, enabling the further discussion of these findings and their implications for practice in the next chapter.

CHAPTER SIX: DISCUSSION AND IMPLICATIONS

The goal of this dissertation was to provide empirical evidence that the use of spatial analysis in enrollment management could provide a benefit for understanding and predicting persistence and graduation. The anticipated outcome was that by combining institutional data with neighborhood data from the U.S. Census, enrollment managers would have an additional resource for predicting a student's probability to persist and graduate. Also, because it utilized existing secondary data, enrollment managers would be able to geographically identify prospective students much earlier in the recruitment process. This identification could be used to recruit students most likely to persist, improving retention and graduation outcomes. Additionally, this identification would enable earlier intervention for students most likely to need assistance and resources, once again improving retention and graduation outcomes.

In this final chapter, I discuss the key findings of this research in relation to each of my research questions. For each, I begin with a discussion of the previously reviewed literature to address whether my findings are supported by this literature. I also will discuss findings in the context of the theoretical frameworks guiding this study, including the forms of capital and social reproduction. Finally, I conclude with implications for practice within enrollment management and directions for future research.

Research Question 1 Discussion.

What are the geospatial patterns of enrollment for the institution of study with respect to: neighborhood household income levels; neighborhood racial/ethnic compositions; neighborhood education levels; and resident and nonresident enrollees?

To answer this first research question, I employed the use of maps to illustrate geospatial patterns of enrollment for The University of Arizona. Using these maps, I then disaggregated by resident and nonresident students, projecting data concerning how these neighborhood patterns are apparent for income levels, ethnic minority population levels, and education levels. The findings of this research indicated spatial analysis can be used as a valuable resource for enrollment management. Since the persistence-related student characteristics can be quantified and mapped, target populations can be spatially identified and subsequently recruited based on institutional goals.

From my review of the literature on spatial analysis within enrollment management, I expected to have two main results from the analysis. The first was that I would see visual patterns rather than random distributions of enrollments (Messner & Anselin, 2004; Robinson, 1959; Mitchell, 2005). The second expectation from the literature was that I would see higher concentrations of enrollments in those neighborhoods nearer the university since "near things are more related than distant things" (Tobler, 1970, p. 236; McConnell, 1965; Kariel, 1968). Given the literature, this was a phenomenon that would be apparent with the states nearest Arizona. Spatial density of enrollments would increase within Arizona, and continue to increase in

Tucson, and finally within neighborhoods nearest the university (McConnell, 1965; Kariel, 1968).

My theoretical framework provided the lens with which to examine or understand in more depth these findings. Examining the spatial distribution of enrollments through the forms of capital and social reproduction allowed this research to explore relationships between class and educational attainment. Measuring class by neighborhood census characteristics, I looked for patterns associated with enrollment and class—were enrollments mostly associated with wealthy, well-educated, and Caucasian neighborhoods? Did predominantly poor, uneducated, ethnic minority neighborhoods produce enrollments? Through social reproduction theory, the expectation would be that an uneducated neighborhood would replicate the lack of education and reduce the probability that a student would enroll at the institution. The privilege of wealth, the privilege of educational attainment and the privilege of being in the majority population would be reproduced, and these areas would evidence a greater number of enrollments.

Spatial distribution of enrollments.

With this understanding of expectations, I discuss my findings regarding the spatial distribution of enrollments from the first series of maps (Figures 3-11). The first finding was that enrolling students did not appear to be randomly distributed. Students were concentrated in selected areas, while nonexistent in other areas. This was suggested by the maps of Continental United States (Figures 3-4), but more apparent on maps beginning with the state levels for California (Figure 5) and Arizona (Figure 6). At these

state levels, the concentrations appeared to be the result of areas with high populations. It is true that the higher population areas of Los Angeles, San Diego, San Francisco, and Sacramento yielded more California enrollees than other cities. Similarly in Arizona, Phoenix and Tucson yielded the greatest number of enrollees. By increasing the scale of the maps to 1:100,000 the differences became clear. There were obvious delineations of enrollment distributions that were not the result of population alone, but were related to median income, race/ethnicity, and education levels of neighborhoods.

Mapping the spatial distribution of enrollments for The University of Arizona also evidenced the geographically extensive reach of this university. By projecting four-years of enrollments onto maps of the Continental United States, it was easy to visualize the distribution. Although scale and stacking influenced perception, the patterns clearly demonstrated that The University of Arizona is a national university. This was supported by the descriptive findings regarding enrollments (Tables 22-23). With the first map illustrating the western extent of the United States (Figure 3), the dot density increased, making it clear that although national in its reach, the university enrolled the majority of students from the western side of the nation—Arizona, California, Washington, and Texas. On the second map of enrollments (Figure 4), it was easy to see concentrations of students originating from the states of Illinois, New York, and New Jersey. This was partially in agreement with the literature on spatial analysis. Western states are closer to the institution and subsequently yielded higher enrollments. However, several neighboring states did not yield substantial numbers of students, and eastern cities with high populations yielded a significantly larger number of students. Once the maps

displayed only Arizona, the literature held true, as Phoenix and Tucson produced significant numbers of students, and the number increased as the university became closer. Enrollments were concentrated in neighborhoods adjacent to the university, but were not as dense in neighborhoods within two miles of the university, as compared to the suburbs of the city. In these areas, enrollments seemed to be influenced by factors other than the proximity to the university. To uncover the nature of these factors, I examined enrollments in the context of three student characteristics associated with the student's neighborhood (block group). I used three series of maps to illustrate distributions enrollments in the context of median household income levels (Figures12-16), education levels (Figures 17-21), and percent of the neighborhood population classified as ethnic minority (Figures 22-26).

Median household income levels.

From the maps showing the distribution of median household income levels (Figures12-16), I discuss three findings. The first was that when viewing the Tucson maps, it was apparent that although high income neighborhoods in the northern portion of Tucson yielded more students, a substantial number of students were originating from the lower-income neighborhoods featured on the two maps. Similarly, the map of Nogales illustrated low-income neighborhoods producing substantial enrollments. In Tucson, distributions from the lowest quintile ($0-$40,000) appeared to be geographically influenced with more coming from this quintile in the northern extent of the map (Figure 12), compared to the southern extent of the map (Figure 13). This may be the result of my choice to use quintiles for disaggregation. Lowest quintile neighborhoods in the

northern extent could potentially be at the upper end of the quintile's scale, while those in the southern extent could be in the lower end of the quintile. It may also be the result of other factors, such as GPA, race/ethnicity, and neighborhood educational levels impact enrollment in ways similar to the way in which they impact retention (as indicated in the statistical analysis associated with Research Question 2). Also, institutional factors not addressed in this research would impact enrollments, persistence, and graduation. When examining findings for median income levels in the other two cities, the results changed dramatically with respect to the stratification of neighborhoods. For Phoenix, low-income neighborhoods yielded few enrollees, while there were definite concentrations of students in the higher income neighborhoods. After having viewed the results for Tucson, the difference was striking. For Los Angeles, this same pattern was present, although more expected since nonresidents were paying $10,000 more a year in tuition, and by default would need to have additional financial resources at their disposal.

Neighborhood education levels.

Using the neighborhood's percentage of males with a bachelor's degree or higher as a proxy for social capital, akin to first generation students, I was able to illustrate distributions of education (Figures 17-21). Compared to maps showing income levels, neighborhoods with similar education levels appear to be more grouped. While enrollments from Tucson and Nogales did not seem to be adversely affected by low neighborhood education levels, the opposite was true for Phoenix and Los Angeles. In these cities, education greatly impacted enrollments. Also, I noted greater delineations in education levels with high and low level neighborhoods clustering. For enrollment

managers, visualizing education levels allows for greater understanding of the data behind enrollment phenomenon.

Percent of neighborhood that is ethnic minority.

By examining enrollments along with projections of ethnic minority neighborhoods, patterns for Phoenix and Los Angeles, show the university does not recruit minorities well outside of Tucson and Nogales. Based on the theoretical frameworks, these neighborhoods would not be expected to produce enrollments, as disadvantage begets disadvantage.

Research Questions 2 and 3: Discussion of Findings

When using geospatial analysis, what are the characteristics associated with a student's probability to persist (or graduate)?

- How do neighborhood household income levels predict persistence (4-year graduation rates)?
- How do neighborhood racial/ethnic compositions predict persistence (4-year graduation rates)?
- How do neighborhood education levels predict persistence (4-year graduation rates)?

Due to similarities within the literature regarding influences on first year persistence and graduation, I discuss the findings related to these two questions together.

This common discussion also enables the contrast of student characteristics important in the first year, yet diminished four years later. As a reminder, this research intentionally does not account for institutional characteristics in the analysis, and any interpretations of the results should be considered within this context. Also, an additional limitation in discussing the results is that findings for RY1were based on 21,838 students over four years of enrollment, while RY4G results are based on 5,043 students from a single enrollment year. In spite of this, I feel the comparison is worthwhile in that it highlights the issue of focusing on year-one retention without understanding what ultimately results in a student's graduation.

How the theoretical framework guides Research Questions 2 and 3.

Interpreting the results of the logistic regressions relating to persistence and graduation through my theoretical framework allowed this research to explore relationships between class and educational attainment. Measuring class by neighborhood census characteristics, I looked for odds ratios positively associated with enrollment and class—were year-one retention and graduation mostly associated with indicators of wealth, education, and race/ethnicity? Through social reproduction theory, the expectation would be that students from disadvantaged neighborhoods would fail to persist at higher rates, and their probability to graduate would be adversely affected.

Student characteristics impacting persistence and graduation.

This research used several student and institutional characteristics to examine persistence and graduation, and a review of each, along with the expectations are

discussed below. Demographic student characteristics included gender and race. Academic student characteristics included GPA and standardized test scores. Geographic characteristics included five location-dependent student neighborhood characteristics: residency (in-state tuition); percent of the neighborhood classified as an urban area; percent of the neighborhood classified as an urban cluster; percent of the neighborhood that is the same as the ethnic minority being analyzed; and the student's distance from the university. Economic capital student characteristics included median household income. Cultural and social capital student characteristics included: the number of tests taken by the student, the timing of the submission of the application for admission, and the percent of males in a neighborhood with a bachelor's degree or higher.

Demographic student characteristics: gender and race/ethnicity.

For the demographic student characteristic of gender, I anticipated a small, or nonexistent effect based on the literature (Adelman, 2006; Alexander, Riordan, Fennessey, & Pallas, 1982). If an effect was present, it would result in females having higher probabilities when compared to males in the sample (Astin, 1993; Kanarek, 1989; Bowen & Bok, 1998; U.S. Department of Education, 2008). In my results, I did not find any evidence that gender influenced the probability of persistence or graduation at The University of Arizona. This was consistent with literature indicating any increases in female graduation rates were often explained by academic history, rather than gender differences (Adelman, 2006).

For race/ethnicity, I expected to find several differences since race has been shown to be an indicator of college readiness that results in higher persistence. This lack of readiness often results in minorities persisting at lower rates (Light & Strayer, 2002; Venti & Wise, 1983; Porter, 1990; Tinto, 1993). In spite of this, other research indicates that when controlling for other factors such as SES and gender, ethnic minority students persisted at higher rates than Caucasian students (Light & Strayer, 2002; Alexander, Riordan, Fennessey, & Pallas, 1982). Since my study controlled for gender, SES and several other student characteristics, I was unsure how race would influence my results. From my findings, it is clear that race has a significant impact upon persistence and graduation at The University of Arizona. When controlling for other factors, an ethnic minority student had a higher probability of persisting to the second year of enrollment (with the exception of Native Americans, as discussed later) (Tables 73-74). However, ethnic minority status became a negative influence on probability of graduation, or persisting past the fourth year. Based on these results the literature, my findings were consistent for both year-one persistence and graduation.

As mentioned, Native American students did not have a higher probability when controlling for other factors. I feel my findings regarding Native American students provide a major contribution to the literature because this is a population that is often difficult to research and produce significant results. In my research, findings for resident and nonresident Native American students show the importance of doing well in high school, as GPA increased the probability to persist and graduate. Additionally, the two proxies for cultural capital, timing of the application submission and number of tests,

increased the probability of both year-one retention and of graduation for resident populations. Perhaps the most telling, and most unfortunate finding was the severely negative effect of a Native American neighborhood on a Native American student with respect to RY1. At .074, a student from a neighborhood that has no Native American population is 13 times more likely to persist in the first year than a student who is from a neighborhood that is 100% (this would include all Native American reservations).

A second major finding related to race, and much more encouraging relates to Hispanic students at The University of Arizona. This research demonstrated that for Hispanic students originating from overwhelmingly Hispanic neighborhoods, there is a higher probability for persistence at the University of Arizona. Ceteris paribus, a Hispanic student from an all Hispanic neighborhood is more than three times as likely to persist to the second year when compared to a Hispanic student coming from an all-Caucasian neighborhood. The Hispanic student also improves their chances by having a higher GPA, and having higher neighborhood education levels, but these student characteristics were not as influential as the percentage of the neighborhood that was Hispanic. Given the university's desire to become a Hispanic serving institution, The University of Arizona has committed substantial resources to assisting this population through first-year retention efforts and programs targeting underserved populations. It is critical to note that the contributions of these efforts are not accounted for in this analysis, as these efforts are post enrollment. Given the success of these students in RY1, it should be encouraging to the institution. However, the effect is not present in the RY4G indicating that additional resources may be necessary in the following years to maintain

the early success. As this is the result from a single institution, it seems advantageous to build upon this research, and analyze the influence of post-enrollment characteristics. Additionally, it would further inform research by analyzing whether this phenomenon is present at other institutions within the state, and at other similar institutions.

Academic student characteristics

The academic student characteristics used in this analysis were high school GPA and standardized test scores. Academic variables have a strong relation to college graduation in the literature (Alexander, Riordan, Fennessey, & Pallas, 1982). In most literature, GPA has proven to be an excellent predictor of both persistence and graduation (Astin, Tsui, & Avalos, 1996; Pike & Saupe, 2002). Given the strong record of prediction in past studies, I expected GPA to have a strong effect in predicting student outcomes at The University of Arizona. For standardized test scores, I was less convinced. Many of the studies reviewed found a relationship between high test scores and persistence (Bowen & Bok, 1998; Vars & Bowen, 1998; Astin, Tsui, & Avalos, 1996) . However, these findings of a positive relationship often noted that it was the combination of high test scores with a high GPA, or institutional characteristics, or some other factors that were the most likely influence on outcomes (Burton & Ramist, 2001; Wilson, 1983; Bowen & Bok, 1998). As my study was controlling for other factors, I did not expect a strong influence from standardized test scores.

The results were as anticipated. While significance was observed, it rarely exerted a strong influence. For residents, test scores were significant for RY4G for Asian

Americans and Caucasians. An Asian American student with the highest possible SAT score could increase their probability to graduate by nearly five times over an Asian American student with the lowest possible score.

Geographic student characteristics

As mentioned earlier in this chapter, five geographic characteristics were included as student and neighborhood characteristics (residency (in-state tuition); percent of the neighborhood classified as an urban area; percent of the neighborhood classified as an urban cluster; percent of the neighborhood that is the same as the ethnic minority being analyzed; and the student's distance from the university). Of these, all were expected to contribute to outcomes for persistence and graduation at varying levels. From the literature, it seemed residency could potentially have a negative influence akin to the potential effect of wealth/poverty. Also, the presence of an existing social structure that residents would be more likely to have in place could affect students in both negative and positive ways (Tinto, 1993). Originating from highly urbanized areas or urbanized clusters was expected to have positive impacts for students. This was based on research on differences between rural and urban students. Expectations were that there could be a small negative effect on persistence outcomes for rural students (Feller, 1974). Percent of neighborhood that is ethnic minority was similar to the previous discussion on race/ethnicity. The expectation would be that an ethnic minority neighborhood could influence outcomes. The final geographic student characteristic was distance from the university. While gravity theory indicated enrollments would be higher in neighborhoods near the university, the literature did not address any relevance to persistence and

graduation. However, given the similarity to Tinto's assertion that existing social networks could impact persistence behaviors, it was expected that gravity theory could influence outcomes in similar manner. For residents, it seemed as though a close proximity to the institution would result in difficulties for selected students, while at the same time greater distances would result in greater expenses for other students, potentially impacting persistence and graduation outcomes.

Residency was used as a categorical variable, and therefore results are aggregated, to control for other student characteristics in this study. When residency was disaggregated, the findings related to residency found GPA and high neighborhood education levels increased the probability to persist at higher levels for residents than nonresidents. The cultural capital variables of number of tests and application in months also improved resident's probability to persist to year two, and to graduate. Nonresidents from neighborhoods with high income levels were more likely to persist to the second year, but this did not impact their graduation probabilities. Instead, graduation was positively influence by GPA, neighborhood education levels, number of tests, timing of the application, and the percent of their neighborhood that was an urban area.

The impact of the three other geographically based student characteristics was surprisingly negligible when disaggregated by race/ethnicity, with the exception African American students. Nonresident African Americans from cities (urban areas) were nearly four times as likely to persist versus African American students who were not from urban areas. In a related finding, resident African Americans from urban clusters (less densely

populated than urban areas) were only 33% as likely to persist as those from other areas. Distance was slightly negative in its effect for RY4G for Asian Americans and Native Americans, but did not seem to influence the persistence of other populations.

Economic capital

For economic capital, this study used median household income level from the census. Poverty has been linked to lower quality schools, unavailability of suitable marriage partners, and reduced exposure to mainstream social networks and conventional role models (Wilson, 1992). However, I felt any effects of low income could be mitigated due to the strong correlation with ethnic minority status. I anticipated positive effects relating to income because as Wilson (1992) noted, it is often what the money is used for that results in advantage (i.e. better schools, cultural activities, etc.).

Findings related to median household income levels were surprising in that they did not have an effect on either year-one retention or graduation outcomes. Most of the literature addressing the influence of affluence is within the overarching category of SES. Since SES is most often defined as some combination of family income and parental education levels, the effect of income was less than expected (Carnevale & Rose, 2004). Based on the findings of this study, the influence of SES on this sample was not due to income levels, but was instead the result neighborhood education levels as discussed below.

Cultural and social capital student characteristics

The social capital variable, the percent of males in a neighborhood with a bachelor's degree or higher, was the most influential. While this influence was significant only for Caucasian and Hispanic students, these categories make up 87% of The University of Arizona's student population. As a proxy to identify first generation students, its identification is important for retention-focused recruitment. The difficulties of first-generation students have been well documented, and this research provides additional support of this research (Pascarella & Terenzini, 1991; Adelman, 2006; Choy, 2001).

For year-one persistence, the effect of neighborhood education levels is substantial for Caucasian and Hispanic populations at the university indicating a student has approximately three times the chance of persisting when compared to enrolling students from neighborhoods with an education level of 0%. While only two race/ethnicity's show this effect, they account for 86% of the university's population in this sample. The effect on graduation probability disappears for Hispanics, but remains a strong predictor of Caucasian students (71% of the sample), both for residents and nonresidents.

Given the opportunity to utilize the institution's detailed dataset, I was able to include a temporal student characteristic not prominent in past literature—the timing of the student's application for admission. As a proxy for cultural capital, it is perhaps a proxy for a student's possession of abilities required to navigate the college experience.

A student possessing such capital would have an advantage by knowing what is rewarded in higher education, and being able to manage personal time more efficiently. Findings from this research support this is rationale, as this student characteristic increased the probability for persistence in four of the five race/ethnicity categories, or 97% of the student population.

With the dataset containing information related to standardized test scores, I was able to access the number of tests each enrolling student submitted. It was surprising that one student had submitted 12 test scores, and that the mean for this category was 1.95 tests submitted. Again, I did not find the number of tests submitted by a student referenced within the literature on the influence of standardized tests scores on outcomes. I defined this as a cultural capital variable based on the assumption that a parent, guardian, or teacher would encourage the student to take multiple tests, or to take a single test multiple times, such as an ACT and SAT, and encourage the student to retake a test that resulted in a lower than desired score. By default a student taking the test earlier would have the opportunity to take it multiple times, and would need to have the means to pay the additional testing fees.

This student characteristic improved a student's probability to graduate for resident Caucasians, Hispanics, and Native Americans. For nonresidents, the effect was not present in analysis of graduation with the exception of Native Americans.

Summary Tables for Findings and Discussion

Summary tables of the findings are presented below and include results for all enrollees disaggregated by residency for year-one retention (Table 73) and graduation (Table 74); disaggregation by race for residents with student characteristics influencing year-one retention (Table 75) and graduation (or persistence after four years) (Table 76); and disaggregation by race for nonresidents with student characteristics influencing year-one retention (Table 77) and graduation (or persistence after four years) (Table 78).

Table 73
Predictors of RY1for all enrollees (N=21,838)

Variable	All	Residents	Nonresident
Race/Ethnicity	—	—	—
Hispanic	1.053	1.151	—
Asian American	1.339**	1.562**	. —
African American	1.134	1.146	—
Native American	.732*	.768	. —
GPA	2.534**	2.720**	2.192**
Percent of bachelor's	2.230**	3.702**	1.705**
Percent of urban area	1.161*	—	1.479**
Number of tests	1.165**	1.160**	1.161**
Application in months	1.085**	1.098	1.060**
Median household income	—	1.000	—
Standardized Test	1.001**	1.001**	—
Constant	0.021**	.011**	.081**

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

Table 74
Predictors of RY4G for all enrollees (N=5,043)

Variable	All	Residents	Nonresident
Race/Ethnicity	—	—	—
Hispanic	.857	.877	—
Asian American	.882	.981	. —
African American	.868	.837	—
Native American	.490*	.490*	. —
GPA	2.590**	2.521**	2.644**
Percent of bachelor's	2.831**	2.629**	2.871**
Percent of urban area	—	—	1.479***
Number of tests	1.164**	1.246**	—
Application in months	1.113**	1.246**	1.094*
Standardized test score	—	1.001**	—
Median household income	—	—	—
Distance from university	1.000	—	1.000
Constant	0.007**	.004**	.024**

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

Table 75
Predictors of RY1, variables for residents, by race/ethnicity (N=21,838)

Variable	Caucasian	Hispanic	Asian American	African American	Native American
Gender	—	—	—	—	—
GPA	2.680**	3.047**	2.629**	2.939**	2.091*
Percent of bachelor's	3.250**	2.790**	—	—	—
Percent of urban area	—	—	—	—	—
Percent of urban cluster	—	—	—	.334*	—
Number of tests	1.140**	1.159	—	—	1.525*
Application in months	1.103**	1.052	1.193**	—	1.118*
Median household income	—	1.000*	—	—	1.000
Standardized test score	1.001**	—	1.002	1.002*	1.001
Distance from university	—	—	—	—	.998
Percent of race/ethnicity (neighborhood)	—	3.228*	—	—	—

* $p < .05$. ** $p < .01$

Table 76
Predictors of RY4G, variables for residents, by race/ethnicity

Variable	Caucasian	Hispanic	Asian American	African American	Native American
Gender	—	—	—	—	—
GPA	2.653**	3.044**	—	2.968	—
Percent of bachelor's	3.088**	—	—	—	—
Percent of urban area	—	—	—	—	—
Percent of urban cluster	—	—	—	—	—
Number of tests	1.188**	1.487**	—	—	—
Application in months	1.143**	1.118	—	—	1.659*
Median household income	—	—	—	—	—
Standardized test score	1.001**	—	1.004**	—	—
Distance from university	—	—	—	—	—
Percent of race/ethnicity (neighborhood)	—	—	.001	—	—

* $p < .05$. ** $p < .01$

Table 77
Predictors of RY1, exp(B), variables for nonresidents by race/ethnicity.

Variable	Caucasian	Hispanic	Asian American	African American	Native American
Gender	—	—	—	—	—
GPA	2.208**	3.210**	—	—	9.345**
Percent of bachelor's	2.101**	2.969	—	—	—
Percent of urban area	1.338**	—	—	3.812	—
Percent of urban cluster	—	—	—	—	—
Number of tests	1.169**	—	—	—	1.880
Application in months	1.090**	—	—	-.132	—
Median household income	—	—	—	—	—
Standardized test score	—	—	—	1.002*	—
Distance from university	—	—	—	—	—
Percent of race/ethnicity (neighborhood)	.492*	—	—	—	.074

* $p < .05$. ** $p < .01$

Table 78
Predictors of RY4G, exp(B), variables for nonresidents by race/ethnicity.

Variable	Caucasian	Hispanic	Asian American	African American	Native American
Gender	—	—	—	—	—
GPA	2.917**	2.932	—	2.968	—
Percent of bachelor's	2.771**	—	—	—	—
Percent of urban area	—	—	—	—	—
Percent of urban cluster	—	—	—	—	—
Number of tests	—	—	—	—	—
Application in months	1.102*	—	1.313	—	—
Median household income	—	—	—	—	—
Standardized test score	—	—	—	—	—
Distance from university	—	—	.999	—	—
Percent race/ethnicity (neighborhood)	—	—	—	—	—

* $p < .05$. ** $p < .01$

Implications for Practice

The management of enrollments requires numerous planning considerations, most notably, what are the goals of the institution? To increase diversity, ensure academic excellence, and improve retention are often at the forefront, but can all goals be accomplished? Through this research into the application of spatial analysis in enrollment management, enrollment managers have an additional resource to target recruitment efforts. Additionally, when the new class arrives on campus, retention efforts can already be underway having identified students who may be at risk, simply by

knowing their neighborhood characteristics. For recruiting those out of state students, enrollment mangers have a resource for understanding the communities from which they recruit. Implications for practice lie in the potential acquisition of knowledge, and the ways in which enrollment mangers are able to use and apply this knowledge for retention-focused outcomes.

Using a retention-focused admissions strategy also could have unintended consequences. As market segmentation and the findings of this study would suggest, an institution blindly implementing a recruitment strategy based solely on pre-enrollment characteristics would become more homogeneous, as more students from neighborhoods with similarly high education levels were enrolled. Continued recruitment from these areas would reproduce homogeneity within the student population resulting in less access and greater stratification. Hypothetically, students from outside the neighborhoods with a higher probability for success would not be recruited (and subsequently enrolled) at the same levels. Although retention would most likely improve and institutional rankings would rise accordingly, the overall diversity of the institution would suffer.

In another implementation, the identified neighborhoods with low levels of education would provide identification of neighborhoods where an institution could focus outreach and recruitment efforts. While the emphasis shifts from retention and resource maximization strategies to achieving greater diversity and reaching underserved populations, this could be beneficial to an institution with related goals.

Directions for Future Research

Along with the potential for using spatial analysis to understand persistence and graduation, future research must consider the limitations associated with this study. The two most notable are the assumptions that census neighborhood designations accurately represent the underlying cultural identity, and that the institutional effects and post enrollment behaviors of the students do not effect outcomes. These limitations were acceptable given the constructs of this research aimed at informing enrollment management; however, the same limitations would not necessarily be appropriate for future studies. To that end, future research on this subject would need to combine the student and neighborhood characteristics with post-enrollment student characteristics and institutional characteristics. Just as the neighborhood of origin was found to influence a student prior to enrollment, so has the university (in a sense, this could also be viewed as a neighborhood) been found to influence students. By combining data representing these two distinct periods (and locations) within a student's development, future research may further the understanding of the intricate nature of institutional characteristics, neighborhood characteristics, student behaviors, and actions that influence persistence and graduation. It is believed that greater understanding of the institution's role in persistence can be elucidated by understanding how the institution interacts with the students in the context of their varied cultural, social, and spatial backgrounds.

For this research, I analyzed six neighborhood characteristics from a census data file containing greater than 5,000 variables. For social science researchers, the potential is intriguing, and with the results of the 2010 U.S. Census to be released in a few years a

tremendous opportunity for discovery will be possible. I believe this forthcoming data set combined with existing analysis tools will allow social science researchers to utilize the data in new ways, providing greater understanding of student behaviors, motivations, and outcomes, while providing greater insight into the institutional forces shaping the college experience.

For enrollment management, spatial analysis provides a resource, an opportunity, and also a responsibility. As the research evolves and predictive capabilities improve, the question already present for enrollment managers is how to use this information to inform student recruitment? If an institution gains the ability to correctly predict graduation outcomes for 90% of enrolling students, does an enrollment manager still take a chance on the one student out of every ten who will graduate in spite of the statistical odds? If the chances are taken only in cases of ethnic minority students, does the Caucasian student lacking cultural capital suffer because they are nearly five times less likely to graduate compared to a similar student possessing high levels of cultural capital? Unfortunately, this research cannot answer these questions, but it can enable enrollment managers to make more informed decisions.

APPENDIX A: VARIABLES ANALYZED FROM U.S. CENSUS

Variables used for analysis of the enrollees include:

- Demographic variables
    - Gender . As 1 for female, 0 for male
    - Race/Ethnicity (see appendix B)
    - High school grade point average
    - Best standardized test score. Calculated using the following formula
    - Residency classification
    - Age at time of enrollment

- Geography-dependent/U.S. Census information
    - Distance from the university (Euclidian)
    - Type of neighborhood, urban/rural (P005)
    - Neighborhood composition by ethnicity
    - Neighborhood composition by educational attainment (P113)
    - Neighborhood composition by median household income in 1999 (P053)
    - Aggregate household income by age of householder (P057)
    - Neighborhood composition by Aggregate public assistance income (P073)
    - Family income by levels (P076)
    - Median family income (P077)
    - Ratio of income to the poverty level (P088)

APPENDIX B: FORMULAS FOR VARIABLES DERIVED FROM U.S. CENSUS

Percent of males who are college graduates
Variable "PercentOfMalesCollegeGrads" = ((P037015 + P037016 + P037017 + P037018)/P037002)

Where P037002 is the total universe of males in the neighborhood and:

- P037015 is the number of those males with a Bachelor's degree as their highest level of academic attainment;
- P037016 is the number of those males with a master's degree as their highest level of academic attainment;
- P037018 is the number of those males with a Doctorate degree as their highest level of academic attainment.

Percent of Population Falling within an Urban Area (Neighborhood)
Variable "PercentOfPopulationUrban" = (P005002/P005001)

Where P005001 is the total universe of residents in a given neighborhood and:

P005002 is the total number of those residents residing in an urban area. Includes those living inside an urbanized area, and those living inside an urban cluster.

WORKS CITED

ABOR. (2005). *The University of Arizona changing directions: Focused excellence.* Retrieved April 1, 2007, from Arizona Board of Regents (ABOR): http://www.asu.edu/changingdirections/papers/uofa.htm

ABOR. (2009, September). *Policy name: Requirements for resident status.* Retrieved March 1, 2010, from Arizona Board of Regents (ABOR): http://www.abor.asu.edu/1_the_regents/policymanual/chap4/4-203.pdf

ACT. (2007). *ACT-SAT concordance tables.* Retrieved May 3, 2009, from ACT: http://www.act.org/aap/concordance/pdf/report.pdf

ACT. (2009). *ACT educational opportunity service (EOS): Results.* Retrieved October 1, 2009, from ACT: http://www.act.org/eos/results.html

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment.* Washington, D.C.: U.S. Government Printing Office.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college.* Washington, D.C.: U.S. Department of Education.

Alexander, K., Riordan, C., Fennessey, J., & Pallas, A. (1982). Social background, academic resources, and college graduation: Recent evidence from the National Longitudinal Survey. *American Journal of Education, 90* (4), 315-333.

America's best colleges. (2009, September). *US News and World Report.*

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

Astin, A. W. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education* (Vol. American Council on Education Series on Higher Education). Washington/New York: American Council on Education and Macmillan.

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

Astin, A. W., & Oseguera, L. (2003). *Degree attainment rates at American colleges and universities.* Los Angeles: Higher Education Research Institute.

Astin, A. W., Oseguera, L., Sax, L. J., & Korn, W. S. (2002). *The American freshman: Thirty-five year trends, 1966-2001.* Los Angeles, California: Higher Education Research Institute, Graduate School of Education and Information Studies, University of California.

Astin, A. W., Tsui, L., & Avalos, J. (1996). *Degree attainment rate at American colleges and universities: Effects of race, gender, and institutional type.* Los Angeles: Higher Education Research Institute, UCLA.

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis.* Essex: Addison Wesley Longman Limited.

Baird, L. (2000). College climate and the Tinto model. In J. Braxton, *Reworking the Student Departure Puzzle* (pp. 62-80). Nashville: Vanderbilt University Press.

Baker, M. J. (Ed.). (1984). *Macmillan dictionary of marketing and advertising.* London: Macmillan.

Bean, J. P. (1990). Why students leave: Insights from research. In D. Hossler, J. P. Bean, & Associates, *The Strategic Management of College Enrollments* (pp. 147-169). San Francisco: Jossey-Bass.

Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research, 55* (4), 485-540.

Becker, G. S. (1964). *Human capital.* Chicago: University of Chicago Press.

Becker, G. S. (2008). *Human capital: The concise encyclopedia of economics.* Retrieved February 16, 2010, from Library of Economics and Liberty: http://www.econlib.org/library/Enc/HumanCapital.html

Berger, J. B. (2000). Optimizing capital, social reproduction, and undergraduate persistence. In J. M. Braxton, *Reworking the Student Departure Puzzle* (pp. 95-125). Nashville: Vanderbilt University Press.

Berger, J. B. (1997). Students' sense of community in residence halls, social integration, and first-year persistence. *Journal of College Student Development, 38* (5), 441-452.

Berger, J. B., & Milem, J. F. (1999). The role of student involvement and perceptions of integration in a causal model of student persistence. *Research in Higher Education, 40* (6), 641-664.

Berkner, L. K., Cuccaro-Alamin, S., & McCormick, A. C. (1996). *Descriptive summary of 1980-1990 beginning postsecondary students: Five years later (NCES 96-155).* National Center for Education Statistics, U.S. Department of Education. Washington, D.C.: U.S. Government Printing Office.

Berkner, L. K., He, S., & Cataldi, E. (2002). *Descriptive summary of 1995-1996 beginning postsecondary students: Six years later.* Washington, D.C.: U.S. Department of Education, National Center for Education statistics.

Berry, J. K. (1987). A mathematical structure for analyzing maps. *Environmental Management, 11* (3), 317-325.

Birnbaum, R. (2000). *Management fads in higher education: Where they come from, what they do, why they fail.* San Francisco: Jossey-Bass Publishers.

Black, J. (Ed.). (2001). *The strategic enrollment management revolution.* Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers.

Blackwell, D. L., & McLaughlin, D. K. (1999). Do rural youth attain their educational goals? *Rural Development Perspectives, 13* (3), 37-43.

Bourdieu, P. (1984). *Distinction.* Cambridge: Harvard University Press.

Bourdieu, P. (1977). *Reproduction in education, society, culture.* Beverly Hills, California: Sage.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). New York: Greenwood Press.

Bourdieu, P., & Passeron, J.-C. (1977). *Reproduction: In education, society and culture.* (R. Nice, Trans.) London: Sage.

Bowen, W. G., & Bok, D. (1998). *The shape of the river.* Princeton, New Jersey: Princeton University Press.

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities.* Princeton: Princeton University Press.

Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: Educational reform and the contradictions of economic life.* New York: Basic Books.

Bowles, S., & Gintis, H. (2002). Schooling in capitalist America revisited. *Sociology of Education, 75,* 1-18.

Bowen, W. M., Salling, M. J., Haynes, K. E., & Cyran, E. J. (1995). Toward environmental justice: Spatial equity in Ohio and Cleveland. *Annals of the Association of American Geographers, 85,* 641-663.

Brack, I., & Martin, D. (1989). The generation of spatial population distributions from census centroid data. *Environment and Planning, 21* (4), 537-543.

Braxton, J. M. (2000). *Reworking the student departure puzzle.* Nashville: Vanderbilt University Press.

Bridgeman, B., Pollack, J., & Burton, N. W. (2008). *Predicting grades in different types of college courses.* New York: College Board Publications.

Burton, N. R., & Ramist, L. (2001). *Predicting success in college: SAT studies of classes graduating since 1980 (College Board report 2001-2).* New York: College Board.

Cabrera, A. F. (1990). Exploring the effects of ability to pay on persistence in college. *The Review of Higher Education, 13* (3), 303-336.

Cabrera, A. F. (1994). Logistic Regression Analysis in Higher Education: An Applied Perspective. In *Higher Education: Handbook of Theory and Research* (Vol. X, pp. 225-256). New York: Agathon Press.

Caison, A. L., & Baker, B. A. (2007). Assessing the impact of differential operationalization of rurality on studies of educational performance and attainment: A cautionary example. *Professional File (106), Winter,* 1-16. Association for Institutional Research.

Camara, W., & Kimmel, E. (2005). *Choosing students: higher education admissions tools for the 21st century.* London: Taylor 7 Francis Group (Routledge).

Carbonaro, W. (1998). A little help from my friend's parents: Intergenerational closure and educational outcomes. *Sociology of Education, 71,* 295-313.

Carnevale, A. P., & Rose, S. J. (2004). Socioeconomic status, race/ethnicity, and selective college admissions. In R. D. Kahlenberg (Ed.), *America's Untapped Resource: Low-Income Students in Higher Education* (pp. 101-156). New York, New York: The Century Foundation.

Carter, D. J., & Wilson, R. (1997). *Minorities in higher education: 1996-1997 fifteenth annual status report.* Washington, D.C.: American Council on Education, Office of Minority Concerns.

Carter, D. (2003). Secondary analysis of data. In F. K. Stage, & K. Manning, *Research in the College Context: Approaches and Methods* (pp. 153-168). New York: Brunner-Routledge.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin, 39* (7), 3-7.

Chickering, A. W., & Reisser, L. (1993). *Education and identity.* San Francisco: Jossey-Bass.

Choy, S. P. (2001). *Students whose parents did not go to college: Postsecondary access, persistence, and attainment.* Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Choy, S. P. (2002). *Access and persistence: Findings from 10 Years of longitudinal research on students.* Washington, D.C.: American Council on Education.

Coleman, J. (1988). Social capital and the creation of human capital. *American Journal of Sociology,* S95-S120.

College Board. (2008). *Understanding your SAT scores.* Retrieved January 13, 2009, from College Board: http://www.collegeboard.com/prod_downloads/student/testing/sat/scores/Understanding-Your-SAT-Scores.pdf

College Board. (2009a). *Recruitment plus.* Retrieved October 1, 2009, from College Board: http://professionals.collegeboard.com/higher-ed/recruitment/recruitment-plus

College Board (2009b). *PSAT/NMSQT: Student search service*. Retrieved December 9, 2009, from College Board: www.collegeboard.com/student/testing/psat/reg/sss.html

Cope, R. G. (1981). *Strategic planning, management, and decision making. AAHE-ERIC/higher education research report, no. 9.* Washington, D.C.: American Association for Higher Education.

Cowton, C. J. (1998). The use of secondary data in business ethics. *Journal of Business Ethics, 17* (4), 423-434.

Creager, J. A. (1968). Use of research results in matching students and colleges. *Journal of College Student Personnel, 9*, 312-319.

Dika, S. L., & Singh, K. (2002). Applications of social capital in educational literature: A critical synthesis. *Review of Educational Research, 72*, 31-60.

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* (April), 189-201.

DiMagio, P., & Mohr, J. (1985). Cultural capital, educational attainment, and marital selection. *American Journal of Sociology, 90* (6), 1231-1261.

Dolence, M. G. (1993). *Strategic enrollment management: Cases from the field.* Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers.

Durkheim, É. W. (1973). *Moral education: A study in the theory and application of the sociology of education.* Detroit: The Free Press.

Ehrenberg, R. (2000). *Tuition rising: Why college costs so much.* Cambridge: Harvard University Press.

Elkins, S. A., Braxton, J. B., & James, G. W. (2000). Tinto's separation stage and its influence on first-semester college student persistence. *Research in Higher Education, 41* (2), 251-268.

Ethington, C. A., Thomas, S. L., & Pike, G. R. (2003). Back to basics: Regression as it should be. In *Higher Education Handbook of Theory and Research* (Vol. 27, pp. 263-293). New York: Agathon Press.

Fact book. (2004). Retrieved May 5, 2009, from Office of Institutional Research and Planning Support: http://oirps.arizona.edu/files/Fact_Book/Factbook04_05.pdf

Fact book. (2007). Retrieved April 1, 2007, from Office of Institutional Research and Planning Support: http://oire.arizona.edu/files/Fact_Book/UA_Factbook06-07.pdf

Feller, B. A. (1974). *Rural-urban differences in scholastic performance in college.* Memphis, Tennessee: Eric Document Reproduction Service, No.ED096026.

Fortune, J. C., & McBee, K. (1984). Consideration and methodology for the preparation of data files. In D. J. Bowering, *Secondary Analysis of Available Data Bases* (pp. 27-49). San Francisco: Jossey Bass.

Fotheringham, A. S., & Rogerson, P. A. (2009). *The Sage handbook of spatial analysis.* Thousand Oaks, California: Sage Publications Inc.

Fotheringham, A. S., Brundson, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships.* London: John Wiley & Sons Ltd.

Frankfort-Nachmias, C., & Nachmias, D. (1992). *Research methods in the social sciences* (4th Edition ed.). London: Edward Arnold.

Freeman, C. E. (2004, November). *Trends in educational equity of girls & women 2004*. Retrieved March 25, 2010, from National Center for Education Statistics: http://nces.ed.gov/pubs2005/2005016.pdf

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

Fuller, W. C., Manski, C. F., & Wise, D. A. (1982, Fall). New evidence on the economic determinants of postsecondary schooling choice. *Journal of Human Resources*, 477-495.

GeoLytics. (2010). *GeoLytics - US census demographic and GIS data products*. Retrieved February 21, 2010, from GeoLytics: http://geolytics.com/USCensus,Census-2000-Long-Form,Products.asp

Gibbs, R. M. (1998). College completion and return migration among rural youth. In R. M. Gibbs, P. L. Swaim, & R. Teixeira, *Rural Education and Training in the New Economy.* Ames, Iowa: Iowa State University Press.

Giroux, H. A. (1981). *Ideology, culture, and the process of schooling.* Philadelphia: Temple University Press.

Gladieux, L. E., & Swail, W. S. (1998). Financial aid is not enough: Improving the odds of college success. *College Board Review, 185*, 16-21, 30-32.

Goldin, C., Katz, L. F., & Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives, 20* (4), 133-156, A1-A4.

Goldrick-Rab, S., & Benson, J. (2006). *Spatial processes in college access: An exploration of Wisconsin (preliminary and incomplete).* Madison, WI: Presented at the

28th Annual Meeting of the Association for Public Policy Analysis and Management, Nov. 2-4, 2006.

Goodchild, M. F., & Janelle, D. G. (2004). Thinking spatially in the social sciences. In M. F. Goodchild, & D. G. Janelle, *Spatially Integrated Social Science* (pp. 3-21). New York: Oxford University Press.

Gratz v. Bollinger, 539 U.S. 244 (United States Supreme Court June 23, 2003).

Hakim, C. (1982). *Secondary analysis in social research: a guide to data sources and methods with examples.* London: Allen and Unwin / Unwin Hyman.

Hamrick, F. A., & Stage, F. K. (2004). College predisposition at high-minority enrollment, low-income schools. *The Review of Higher Education, 27* (2), 151-168.

Hearn, J. (1990). Pathways to attendance at the elite colleges. In P. Kingston, & J. Smart, *Elite Undergraduate Institutions.* New York: Agathon Press.

Hearn, J. (1991). Academic and nonacademic influences on the college destinations of 1980 high school graduates. *Sociology of Education, 64*, 158-171.

Hezlett, S., Kuncel, N., Vey, M., Ahart, A., Ones, D., Campbell, J., et al. (2001). The effectiveness of the SAT in predicting success early and late in college: A meta analysis. Seattle, Washington: Paper presented at the AERA/NCME Annual Meeting.

Holmstrand, P. (2006). *Batchgeocode*. Retrieved June 15, 2009, from batchgeocode.com: http://www.batchgeocode.com/

Horn, L. (1998). *Stopouts or stayouts? Undergraduates who leave college in their first year.* Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: Wiley.

Hossler, D. (1984). *Enrollment management: An integrated approach.* New York: College Entrance Examination Board.

Hossler, D. (1999). Effective admissions recruitment. *New Directions for Higher Education, 1999* (108), 15-30.

Hossler, D. (2000). Effective admissions recruitment. In G. H. Gaither, *Promising Practices in Recruitment, Remediation, and Retention: New Directions for Higher Education, No. 108.* San Francisco: Jossey-Bass.

Hossler, D. (2007). *Fostering student success in the campus community* (First ed.). (G. A. Kramer, Ed.) San Francisco, CA: Jossey-Bass.

Hossler, D., & Anderson, D. K. (2004). The enrollment management process. In M. L. Upcraft, J. N. Gardner, & B. O. Barefoot (Eds.), *Challenging and Supporting the First-Year Student: A Handbook for Improving the First year of College.* San Francisco: Jossey Bass.

Hossler, D., & Bean, J. P. (1990). *The strategic management of college enrollments.* San Francisco: Jossey-Bass.

Hossler, D., Braxton, J., & Coopersmith, G. (1989). Understanding student college choice. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. 5, pp. 231-288). New York: Agathon Press.

Hossler, D., Schmit, J., & Vesper, N. (1999). *Going to college: How social, economic, and educational factors influence the decisions students make.* Baltimore: Johns Hopkins University Press.

Hurtado, S., & Carter, D. F. (1997). Effects of college transition and perceptions of the campus racial climate on Latino college student's sense of belonging. *Sociology of Education, 70* (4), 324-345.

Hurtado, S., Carter, D. F., & Spuler, A. J. (1996). Latino student transition to college: Assessing difficulties and factors in successful adjustment. *Research in Higher Education, 37*, 135-157.

Hurtado, S., Inkelas, K. K., Briggs, C., & Rhee, B. S. (1997). Differences in college access and choice among racial/ethnic groups: Identifying continuing barriers. *Research in Higher Education, 38*, 43-75.

Hutchinson, S. R., & Lovell, C. D. (2004). A Review of methodological characteristic of research published in key journals in higher education: Implications for graduate research training. *Research in Higher Education, 45* (4), 383-403.

Hurtado, S., Milem, J. F., Clayton-Pederson, A., & Allen, W. (1996). *Improving the climate for diversity in higher education institutions (Final report to the Lilly Endowment).* Nashville, TN: Peabody College, Vanderbilt University.

Ihlanfeldt, W. (1980). *Achieving optimal enrollments and tuition revenues.* San Francisco: Jossey-Bass Inc., Publishers.

Ingersoll, R. J. (1988). *The enrollment problem: Proven management techniques.* New York: American Council on Education/Macmillan Publishing Company.

Ishanti, T. T. (2006). Studying attrition and degree completion behavior among first-generation college students in the United States. *Journal of Higher Education, 77* (5), 861-885.

Ishitani, T. T. (2008). How to explore timing of intervention for students at risk of departure. *New Directions for Institutional Research* (137), 105-122.

Ishitani, T. T., & DesJardins, S. L. (2002). A longitudinal investigation of dropout from college in the United States. *Journal of College Student Retention: Research Theory and Practice, 4* (2), 173-201.

Jardine, D. D., & Teodorescu, D. (2003). An introduction to GIS: Concepts, tools, data sources, and types of analysis. *New Directions for Institutional Research, 2003* (120), 5-13.

Johnson, H. G. (1960). The political economy of opulence. *Canadian Journal of Economics and Political Science, 26*, 552-64.

Kanarek, E. A. (1989, April 30-May3). Exploring the murky world of admissions predictions. Baltimore, Maryland: Paper presented at the Twenty-Ninth Annual Forum of the Association for Institutional Research.

Kane, T. J. (1998). Racial and ethnic preferences in college admissions. In C. Jencks, & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 431-456). Washington, D.C.: The Brookings Institution.

Kariel, H. G. (1968). Student enrollment and student interaction. *Annal of Regional Science, 2* (2), 114-125.

Kemerer, F. R., Baldridge, J. V., & Green, K. (1982). *Strategies for effective enrollment management.* Washington, D.C.: American Association of State Colleges and Universities.

King, J. E. (2000). *Gender equity in higher education: Are male students at a disadvantage?* Washington, D.C.: American Council of Education, Center for Policy Analysis.

Kojaku, L. K., & Nunez, A. M. (1998). *Descriptive summary of 1995-1996 beginning postsecondary students with profiles of students entering 2- and 4- year institutions (NCES 98-082).* Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Kotler, P. (1976). Applying marketing theory to college admissions. In *A Role for Marketing in College Admissions* (pp. 54-72). New York: College Entrance Examination Board.

Kotler, P. (1984). *Marketing management: Analysis, planning and control* (Fifth ed.). Englewood Cliffs, New Jersey: Prentice-Hall.

Krivoruchko, K., Gotway, C. A., & Zhigimont, A. (2003). Statistical tools for regional data analysis using GIS. *GIS '03.* New Orleans: ESRI.

Kroc, R., & Hansen, G. (2003). Enrollment management. In W. Knight. Tallahassee, FL: Association for Institutional Research.

Kuh, G. D., & Love, P. G. (2000). A cultural perspective on student departure. In J. Braxton, *Reworking the Student Departure Puzzle* (pp. 196-234). Nashville: Vanderbilt University Press.

Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). *What matters to student success: A review of the literature.* Washington, D.C.: National Postsecondary Education Cooperative.

Kuh, G. D., Schuh, J. H., Whitt, E. J., & Associates. (1991). *Involving colleges: Successful approaches to fostering student learning and personal development outside the classroom.* San Francisco: Jossey-Bass.

Lamont, M., & Lareau, A. (1988). Cultural capital: Allusions, gaps and glissandos in recent theoretical developments. *Sociological Theory, 6* (2), 153-168.

Lareau, A., & Horvat, E. M. (1999). Moments of social inclusion and exclusion race, class, and cultural capital in family-school relationships. *Sociology of Education, 72* (1), 37-53.

Lareau, A., & Weininger, E. B. (2003). Cultural capital in educational research: A critical assessment. *Theory and Society, 32* (5/6), 567-606.

Laskar, J. (1995). The chaotic motion of the solar system. In J. T. Van Tran, *Chaos and Complexity* (pp. 53-62). Paris: Editions Frontieres.

Lee, V., Mackie-Lewis, C., & Marks, H. (2003). Persistence to the baccalaureate degree for students who transfer from community college. *American Journal of Education, 102*, 80-114.

Light, A., & Strayer, W. (2002). From Bakke to Hopwood: Does race affect college attendance and completion? *The Review of Economics and Statistics, 84* (1), 34-44.

Lin, N. (1999). Building a network theory of social capital. *Connections (INSNA), 22* (1), 28-51.

Litten, L. H., Sullivan, D., & Brodigan, D. L. (1983). *Applying market research in college admissions.* New York: College Entrance Examination Board.

London, H. (1992). Transformations: Cultural challenges faced by first-generation students. In L. Zwerling, & H. London, *First Generation Students; Confronting the Cultural Issues* (pp. 5-11). San Francisco: Jossey Bass.

Longden, B. (2004). Interpreting student early departure from higher education through the lens of cultural capital. *Tertiary Education and Management, 10*, 121-138.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2005). *Geographical information systems and science* (2nd Edition ed.). London: John Wiley & Sons Ltd.

Madigan, T. J. (2002). Cultural capital. In D. L. Levinson, P. W. Cookson, & A. R. Sadovnik, *Education And Sociology: An Encyclopedia* (pp. 121-124). New York: RoutledgeFalmer.

Marble, D. F., Mora, V. J., & Granados, M. (1997). Applying GIS technology and geodemographics to college and university admissions planning: Some results from The Ohio State University. *Proceedings of the 17th Annual ESRI User Conference.* San Diego, CA.

Marble, D. F., Mora, V. J., & Herries, J. P. (1995). Applying GIS technology to the freshman admissions process at a large university. *Proceedings of the 15th Annual ESRI Users Conference.* San Diego, CA.

Martin, D. (2009). The role of GIS. In A. S. Fotheringham, & P. A. Rogerson, *The SAGE Handbook of Spatial Analysis* (pp. 25-39). Thousand Oaks, CA: SAGE Publications Inc.

Martin, R. N. (2001). Identifying target enrollment areas to improve diversity. *Proceedings of the 21st Annual ESRI International Users Conference.* Redlands, California: ESRI.

Martin, R. N. (2002). Identifying target enrollment areas to improve diversity, part 2. *Proceedings of the 22nd Annual ESRI International Users Conference.* Redlands, California: ESRI.

Marx, K. (1933, 1849). *Wage-labour and capital.* (L. McCallister, & C. S. Fischer, Eds.) New York: International Publishers Co.

Massey, D. S., (1996). The age of extremes: Concentrated affluence and poverty in the twenty-first century. *Demographics, 33,* 395-412.

Massey, D. S., Gross, A. B., & Shibuya, K. (1994). Migration, segregation, and the geographic concentration of poverty. *American Sociological Review, 59,* 425-445.

Mathiasen, R. L. (1984). Predicting college academic achievement: A research review. *College Student Journal, 18,* 380-386.

McConnell, H. (1965). Spatial variability of college enrollment as a function of migration potential. *The Professional Geographer, 17* (6), 29-37.

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

McDonough, P. M. (1997). *Choosing colleges: How social class and schools structure opportunity.* Albany: State University of New York Press, Albany.

Menard, S. (1995). *Applied logistic regression analysis.* Thousand Oaks: Sage Publications Inc.

Messner, S. F., & Anselin, L. (2004). Spatial analyses of homicide with areal data. In M. F. Goodchild, & D. G. Janelle, *Spatially Integrated Social Science* (pp. 127-144). New York: Oxford University Press.

Milem, J., & Berger, J. (1997). A modified model of college student persistence: The relationship between Astin's theory of involvement and Tinto's theory of student departure. *Journal of College Student Development, 38* (4), 387-400.

Miller, T., Bender, B., Schuh, J., & Associates. (2005). *Promoting reasonable expectations: Aligning student and institutional views of the college experience.* San Francisco: Jossey-Bass.

Mitchell, A. (1999). *The ESRI guide to GIS analysis, Volume 1: Geographic patterns & relationships.* Redlands, California: ESRI Press.

Mitchell, A. (2005). *The ESRI guide to GIS analysis, Volume 2: Spatial measurements & statistics.* Redlands, California: ESRI Press.

Monmonier, M. (1996). *How to Lie with Maps.* Chicago: The University of Chicago Press.

Mora, V. J. (2003). Applications of GIS in admissions and targeting recruiting efforts. *New Directions for Institutional Research, 2003* (120), 15-21.

Morrow, V. (1999). Conceptualising social capital in relation to the well-being of children and young people: A critical review. *Sociological Review, 47,* 744-765.

Mouw, J., & Khanna, R. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal, 27,* 328-336.

Murdock, T. (. (1986). Dissertation: The effect of financial aid on persistence in American higher education. Tucson, Arizona: The University of Arizona.

Musoba, G., & Baez, B. (2009). The cultural capital of cultural and social capital. (J. Smart, Ed.) *Higher Education: Handbook of Theory and Research,* 151-182.

National Commission on the Cost of Higher Education. (1998). *Straight talk about college costs and prices.* Phoenix: Oryx Press.

NCHEMS. (2009, February 18). *Retention rates - First-time college freshmen returning their second year (NCES, IPEDS).* Retrieved February 18, 2009, from The National Center for Higher Education Management Systems (NCHEMS) Information Center: http://www.higheredinfo.org/datamaps.php

Noble, J., Davenport, M., Schiel, J., & Pommerich, M. (1999, July). *Relationships between the noncognitive characteristics, high school course work and grades, and test scores for ACT-tested students.* Retrieved May 24, 2009, from ACT: http://www.act.org/research/reports/pdf/ACT_RR99-04.pdf

Noel-Levitz. (2009). *ForecastPlus for recruitment overview.* Retrieved October 1, 2009, from Noel-Levitz: https://www.noellevitz.com/Our+Services/Recruitment/ForecastPlus/

Nora, A., Attinasi, L. C., & Matonak, A. (1990). Testing qualitative indicators of precollege factors in Tinto's attrition model: A community college student population. *Review of Higher Education, 13* (3), 337-355.

Nuernberger, A. (2001). *GIS classics.* Retrieved January 25, 2008, from Center for Spatially Integrated Social Science: http://www.csiss.org/classics/content/106

OIRPS. (2008). *New freshmen application admission enrollment profile by residency*. Retrieved May 5, 2009, from Office of Institutional Research and Planning Support (OIRPS): http://oirps.arizona.edu/files/Student_Demo/New_Freshmen_Application_Admission_Enrollment_Profile_by_Residency.pdf

O'Quinn, M. D. (1999). Getting above our raising: A case study of women from the Coalfields of southwest Virginia and eastern Kentucky. *Journal of Research in Rural Education*, *15* (3), 181-189.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis, 27*, 286-306.

Ousley, C. A. (2007, November). Can growth, diversity, and academic excellence coexist as enrollment management goals? *Presented at the Association for the Study of Higher Education National Conference*. Louisville, Kentucky.

Pace, C. R. (1980). Assessing diversity among campus groups. In L. L. Baird, R. T. Hartnett, & Associates, *Understanding Student and Faculty Life.* San Francisco: Jossey-Bass Publishers.

Painter, P., & Painter, N. (1982). Placing students for stability and success. In W. R. Lowery, & Associates, *College Admissions Counseling.* San Francisco: Jossey-Bass Publishers.

Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman Persistence and voluntary dropout decisions from a theoretical model. *Journal of Higher Education*, *51* (1), 60-75.

Pascarella, E. T., & Terenzini, P. T. (1998). Studying college students in the 21st century: Meeting new challenges. *The Review of Higher Education*, *21* (2), 151-165.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research.* San Francisco: Jossey-Bass Inc.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (2nd ed.). San Francisco, CA: Jossey-Bass.

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., & Terenzini, P. T. (2004, May - June). First-Generation College Students: Additional Evidence on College Experiences and Outcomes. *The Journal of Higher Education*, 249-284.

Paulsen, M. B., & St. John, E. P. (2002). Social class and college costs: Examining the financial nexus between college choice and persistence. *The Journal of Higher Education, 73,* 189-236.

Pedhazur, E. J. (1982). *Multiple regressions in behavioral research: Explanation and prediction.* New York: Holt, Rinehart & Winston.

Peng, C.-Y. J., So, T.-S. H., Stage, F. K., & St. John, E. P. (2002). The use and interpretation of logistic regression in higher education journals: 1988-1999. *Research in Higher Education, 43* (3), 259-293.

Perna, L. W., & Titus, M. A. (2005). The relationship between parental involvement as social capital and college enrollment: An examination of racial/ethnic group differences. *The Journal of Higher Education, 76* (5), 485-518.

Peters, A., & MacDonald, H. (2004). *Unlocking the census with GIS.* Redlands, California: ESRI.

Pike, G. R., & Kuh, G. D. (2005). First- and second-generation college students: A comparison of their engagement and intellectual development. *Journal of Higher Education, 76* (3), 276-300.

Pike, G., & Saupe, J. L. (2002). Does High School Matter? An Analysis of Three Methods of Predicting First-Year Grades. *Research in Higher Education, 43* (2), 187-207.

Porter, O. F. (1990). *Undergraduate completion and persistence at four-year colleges and universities.* Washington D.C.: National Institute of Independent Colleges and Universities.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology, 24,* 1-24.

Ptolemy, C. (1932). *Geographica.* (E. Stevenson, Trans.) New York: New York Public Library.

Putnam, H. The ‘corroboration’ of theories. In R. Boyd, P. Gasper, & J. Trout, *The Philosophy of Science* (pp. 121-138).

Quirk, M. (2005, November). The best class money can buy: The rise of the enrollment manager and the cutthroat quest for competitive advantage. The secret weapon: Financial aid leveraging. *The Atlantic Monthly*.

Robinson, E. (1959). *An introduction to infinitely many variates.* New York: Hafner.

Rogerson, P. A. (2009). Monitoring changes in spatial patterns. In A. S. Fotheringham, & P. A. Rogerson, *The SAGE Handbook of Spatial Analysis* (pp. 343-354). Thousand Oaks, CA: SAGE Publications Inc.

Salisbury, R. H. (1969). An exchange theory of interest groups. *Midwest Journal of Political Science, 13* (1), 1-32.

Saupe, J., Smith, T., & Xin, W. (1999, May-June). Institutional and student characteristics in student success: First-term GPA, one-year retention and six-year graduation. Seattle, Washington: Paper presented at the Association for Institutional Research.

Schultz, T. (1961). Investment in human capital. *American Economic Review, 51*, 1-17.

Schumacher, S. (2007, January). *Probabilistic versus deterministic data matching: Making an accurate decision.* Retrieved June 15, 2009, from Information Management: http://www.information-management.com/specialreports/20070118/1071712-1.html

Slaughter, S., & Rhoades, G. (2004). *Academic capitalism and the new economy: Markets, state, and higher education.* Baltimore: The Johns Hopkins University Press.

Smedly, B. D., Myers, H. F., & Harrell, S. P. (1993). Minority-status stresses and the college adjustment of ethnic minority freshmen. *Journal of Higher Education, 64*, 434-452.

Smith, B. W., Spinelli, J. G., & Zhou, Y. (2002). Geographic patterns of student enrollment in Ohio's state-assisted universities. *Ohio Journal of Science, 102* (3), 34-39.

Snow, J. (1936). *Snow on cholera.* (B. Richardson, Ed.) New York: The Commonwealth Fund.

Spady, W. (1971). Dropouts from higher education: Toward an empirical mode. *Interchange, 2*, 38-62.

SPSS. (2002). *SPSS 12.5 users manual.* Chicago.

St. John, E. P., Cabrera, A. F., Nora, A., & Asker, E. H. (2000). Economic influences on persistence reconsidered: How can finance research inform the reconceptualization of persistence models? In J. M. Braxton, *Rethinking the Student Departure Puzzle* (pp. 29-47). Nashville: Vanderbilt University Press.

Stampen, J. O., & Cabrera, A. F. (1986). Exploring the effects of student aid on attrition. *Journal of Student Financial Aid, 16*, 28-37.

Stampen, J. O., & Cabrera, A. F. (1988). Is the student aid system achieving its objectives? Evidence on targeting and attrition. *Economics of Education Review, 7*.

Steinberg, S. J., & Steinberg, S. L. (2006). *Geographic information systems for the social sciences: Investigating space and place.* Thousand Oaks, California: Sage Publications, Inc.

Stevens, M. L. (2007). *Creating a class: College admissions and the education of elites.* Cambridge, Mass: Harvard University Press.

Stewart, D., & Kamins, M. A. (1993). *Secondary research: Information sources and methods* (2nd ed.). Newberry Park, California: Sage.

Stewart, E. B., Stewart, E. A., & Simons, R. L. (2007). The effect of neighborhood context on the college aspirations of African American adolescents. *American Educational Research Journal, 44* (4), 896-919.

Stewart, J. Q. (1941). The 'gravitation,' or geographical drawing power of a college. *Bulletin of the American Association of University Professors, 27*, 70-75.

Sullivan, D. F., & Litten, L. H. (1976). Using research in analyzing student markets: A case study. In *A Role for Marketing in College Admissions* (pp. 86-113). New York: College Entrance Examination Board.

Swartz, D. (2002). Social reproduction. In D. L. Levinson, P. W. Cookson, & A. R. Sadovnik, *Education and Sociology: An Encyclopedia* (pp. 551-557). New York: RoutledgeFalmer.

Tate, W. F. (2008). 2008 presidential address: Geography of opportunity: Poverty, place, and educational outcomes. *Educational Researcher, 37* (7), 397-411.

Thomas, G. E. (1981). *Black students in higher education: Conditions and experiences in the 1970s.* Westport, CT: Greenwood Press.

Tierney, W. G. (1992). An anthropological analysis of student participation in college. *The Journal of Higher Education, 63* (6), 603-618.

Tinto, V. (1982). Limits of theory and practice in student retention. *Journal of Higher Education, 36* (5).

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition* (2nd Edition ed.). Chicago: University of Chicago Press.

Titus, M. A. (2006). Understanding college degree completion of students with low socioeconomic status: The influence of the institutional financial context. *Research in Higher Education, 47* (4), 296-317.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*, 234-40.

Trow, M. A. (1984). The analysis of status. In B. W. Clark (Ed.), *Perspectives on Higher Education: Eight Disciplinary and Comparative Views* (pp. 132-164). Berkeley: University of California Press.

Trow, M. A. (2000, March 1). *From mass higher education to universal access: The American advantage.* Retrieved February 10, 2010, from UC Berkeley: Center for Studies in Higher Education (eScholarship): http://escholarship.org/uc/item/9f02k0d1

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd Edition ed.). Cheshire, Connecticut: Graphics Press.

Turley, R. L. (2003). When do neighborhoods matter? The role of race and neighborhood peers. *Social Science Research, 32*, 61-79.

Tynan, A. C., & Drayton, J. (1987). Marketing segmentation. *Journal of Marketing Management, 2* (3), 301-335.

U.S. Census Bureau. (2000). *Appendix A. Census 2000 geographic terms and concepts*. Retrieved April 6, 2009, from U.S. Census Bureau: http://www.census.gov/geo/www/tiger/glossry2.pdf

U. S. Census Bureau. (2002). *Census 2000 urban and rural classification*. Retrieved April 6, 2009, from U.S. Census Bureau: http://www.census.gov/geo/www/ua/ua_2k.html

U.S. Census Bureau. (2003a). *2000 census of population and housing, summary file 3, technical documentation, SF3/11(RV).* Washington, D.C.: U.S. Department of Commerce, U.S. Census Bureau.

U.S. Census Bureau. (2003b). *Technical assessment of ACE revision II.* Washington, D.C.: U.S. Department of Commerce, U.S. Census Bureau.

U.S. Census Bureau. (2006). *Press releases*. Retrieved March 31, 2007, from U.S. Census Bureau: http://www.census.gov/Press-Release/www/releases/archives/population/007910.html

U.S. Census Bureau. (2007a). *2010 census: 2010 census is different*. Retrieved January 4, 2007, from U.S. Census Bureau: http://www.census.gov/2010census/about_2010_census/007622.html

U.S. Census Bureau. (2007b). *U.S. Census Bureau.* Retrieved March 29, 2010, from Census.gov: http://www.census.gov/Press-Release/www/releases/archives/population/009865.html

U.S. Census Bureau. Population Division, (2006). *Table 1: Annual estimates of the population of metropolitan and micropolitan statistical areas: April 1, 2000 to July 1, 2005 (CBSA-EST2005-01).* Retrieved 1 14, 2009, from US Census Bureau: http://www.census.gov/popest/metro/tables/2005/cbsa-01-fmt.xls

U.S. Code. (2007, January 3). *20 USC chapter 28-Higher education resources and student assistance.* Retrieved May 16, 2009, from Office of the Law Revision Counsel, U.S. House of Representatives: http://uscode.house.gov/download/pls/20C28.txt

U.S. Department of Education. (2008). *Digest of education statistics, 2008.* Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences.

USDOE. (2006). *A test of leadership: Charting the future of U.S. higher education.* U.S. Department of Education. Washington, D.C.: Secretary of Education's Commission on the Future of Higher Education.

Valadez, J. (1993). Cultural capital and its impact on the aspirations of nontraditional community college students. *Community College Review, 21* (3), 30-43.

Vars, F. E., & Bowen, W. G. (1998). Scholastic aptitude test scores, race, and academic performance at selective colleges and universities. In C. Jencks, & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 457-479). Washington, D.C.: The Brookings Institution.

Venti, S. F., & Wise, D. A. (1983, June). Individual attributes and self selection in higher education. *Journal of Public Economics*, 1-31.

Wald, A. (1944). On a statistical problem arising in the classification of an individual into one of two groups. *The Annals of Mathematical Statistics*, 145-162.

Walpole, M. (2007). Economically and educationally challenged students in higher education: Access to outcomes. *ASHE Higher Education Report, 33* (3).

Wang, F. (2006). *Quantitative methods and applications in GIS.* Boca Raton, Florida: Taylor & Francis Group.

Weber, M. (1968). Bureaucracy. In G. Roth, & R. Wittich (Eds.), *Economy and Society* (Vols. 1, Conceptual Exposition, pp. 965-1005). New York: Bedminister Press.

Williams, R. W. (1999). The contested terrain of environmental justice research: community as unit of analysis. *The Social Science Journal, 36* (2), 313-328.

Williams, T. E. (1984). Recruiting graduates: understanding student-institution fit. In D. Hossler, *Enrollment Management: An Integrated Approach* (pp. 69-89). New York: College Entrance Examination Board.

Wilson, K. M. (1983). *A review of research on the prediction of academic performance after the freshman year (College Board Report No. 83-2).* New York: College Board.

Wilson, W. (1987). *The truly disadvantaged.* Chicago: University of Chicago Press.

Wilson, W. (1992). Another look at the truly disadvantaged. *Political Science Quarterly*, 639-656.

Wirt, J., Choy, S., Rooney, P., Hussar, W., Provasnik, S., & Hampden-Thompson, G. (2005, June). *The condition of education 2005.* Retrieved March 25, 2010, from National Center for Education Statistics: http://nces.ed.gov/pubs2005/2005094.pdf

Wolfe, A. B. (1932). Population censuses before 1790. *Journal of the American Statistical Association, 27* (180), 357-370.

Wulder, M., & Boots, B. (1998). Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. *International Journal of Remote Sensing, 19*, 2223-2331.

Zhou, Y., & Wu, J. (2005). Modeling university enrollments with ArcGIS. *Proceedings of the 25th Annual ESRI International Users Conference* (pp. 1-9). Redlands, California: ESRI.

Zimmerman, R. (1994). Issues of classification in environmental equity: How we manage is how we measure. *Fordham Urban Law Journal, 21*, 633-669.

Zumeta, W. (2000). Public university accountability to the state in the late twentieth century: Time for rethinking. In R. Weissberg, *Democracy in the Academy* (pp. 93-122). Hauppauge, New York: Nova Science Publishers.

Zweigenhaft, R. L. (1993). Prep school and public school graduates of Harvard: A longitudinal study of the accumulation of social and cultural capital. *The Journal of Higher Education, 64,* 211-225.

Zwick, R., & Sklar, J. G. (2005). Predicting college grades and degree completion using high school grades and SAT scores: The role of student ethnicity and first language. *American Educational Research Journal, 42* (3), 439-464.