Predicting Enrollment of New Criminal Justice Doctoral Programs

William E. Stone[*]

School of Criminal Justice, Texas State University San Marcos, Texas, USA
*Corresponding author: ws03@txstate.edu

Abstract The study reports the results of an attempt to predict enrollment for a newly proposed doctoral program in criminology/criminal justice. The methodology used to create the enrollment projection was a differential equation utilizing a combination of survey data and existing archival data. The study compares the projection results to the first five years of actual enrollment in the program to validate the projection. While the enrollment projection was somewhat off in the first two years, in years three through five the projection was very successful. While this study focuses on a specific program, the methodology was successful and should be applicable to predicting enrollment in a wide range of programs where preexisting populations are not available to form a projection base.

Keywords: population projection, doctoral programs, program justification

Cite This Article: William E. Stone, "Predicting Enrollment of New Criminal Justice Doctoral Programs." *American Journal of Educational Research*, vol. 3, no. 10 (2015): 1208-1215. doi: 10.12691/education-3-10-1.

1. Introduction

It is clear that very few universities or university systems can continue to expand overall, at the same rate they have done in the past. For many universities, space and financial limitation, have caused administrators to reconsider the viability of overall growth and to change their focus to "targeted" growth. By targeted growth, we mean allowing those areas of the university to grow that represent a special community, social, or financial need for the university. The obvious corollary to this approach is controlling or limiting the areas that don't represent a special need [1].

In higher education, there is much greater economic competition than there is in the traditional public schools. The student population can, within reason, choose where to spend their educational dollars. This simply means that university administrators are forced to consider the financial viability of educational programs in general and especially, new proposed degree plans. The author is not proposing or supporting a strict financial model of educational growth. There are clearly other important needs but university administrators seem to give priority to financial considerations over social, instructional, academic or other more abstract needs.

This phenomena is especially true when considering new doctoral programs regardless of the specific program major. By their nature, doctoral programs are very expensive because of small class size, high faculty cost, and sometimes exorbitant startup equipment costs. When a new doctoral program is proposed, a "viability plan" is normally required. While the plan must consider faculty credentials, library resources and all the other traditional academic resources necessary for a quality program, much of the upper university administration is likely to focus on the financial viability of the program. Questions such as "program need" (how many students will be enrolled) become a very central part of the new degree program proposal. This study reports the results of just such a population viability study for the establishment of a new doctoral program in criminal justice.

The need to project enrollment is not new to the educational profession. Attempts at educational population projection can clearly be traced back to the first part of the twentieth century. Most of the early population prediction models were developed for large public school systems which had an obvious need to predict future enrollment for new construction and financial planning purposes [2]. Large public school systems represented an environment where; reliable archival data sets, research expertise, need (public pressure), and an environment conducive to the application of predictive modeling techniques all come together. For example, if we are trying to predict a school systems enrollment population of 1st graders, we would need; reasonably accurate demographic data sets on the number of children of the correct age, the promotion/failure rate in the first grade from the previous year, and the voluntary transfer rate to and from other schools from the previous year. These might be called the "core independent variables" which would account for most of the variance in the dependent variable, the number of enrolled 1st graders, and would be reasonably available in most large urban areas. Nothing improves the ability to make good predictions like large quality data sets on most of your independent variables that can be used for regression models or other similar equations [3].

This is not an attempt to denigrate the early prediction efforts. Many early studies were very sophisticated and

they addressed many subtle independent variables not included above. The point instead, is to illustrate that when trying to model the population of a new doctoral program most of the environmental factors are not conducive to the application of our more powerful prediction techniques. The "core independent variables" impacting enrollment for most new doctorial programs are "poorly known" or unknown, and the available archival data sets are of questionable quality. This means the most common multivariate research tools are of very limited applicability [4].

We traditionally use multivariate statistics to analyze the past and then use the trend to predict the future. For the example in question, new doctoral programs, there is no past data to utilize. To project populations under these restraints the researcher must be willing to; reexamine the fundamentals of educational modeling, carefully study the research environment, consider less sophisticated mathematical prediction techniques, and accept some educated/professional guessing in the projection model. This approach was well documented by Weisman in his 1994 article *Enrollment Projections: Combining Statistics and Gut Feelings* [5].

1.1. Types of Educational Models

According to the early work of Correa, educational population prediction models identified in the literature can be classified into three basic types; micro-models, macro-models and hybrid models (hybrid models also include most simulation models) [6].

1.1.1. Micro-Educational Models

The first, micro-educational models are relate to the educational process itself. They describe and predict behavior internal to the learning process such as the interaction between teachers and students, students and administrative procedures, or between students themselves. Early examples of such models are the associationist model developed by Bush and Mosteller [7], as well as early learning tree models such as those by Restle [8] and Scandura [9]. This type of model would be useful in examining the population impact of programmatic variables and instructional methodologies in a new doctoral program. They would allow you, to some degree, predict the population impact of; required course sequences, online courses vs traditional courses, or a student cohort structure on student retention. Not surprisingly, much of the micro based research has been done on high attrition groups like college freshmen, because of the large available data sets, and the high financial impact of freshman attrition.

These models will not assist in predicting new student admissions but will instead be useful in examining attrition and its impact on populations past the initial admission point. This will make a significant contribution to the viability of program enrollment predictions which normally require projections for an initial five to ten year period. When examining enrollment populations, it must always be considered that "current enrolled population" is a group made up of initial enrollment, plus new students transferring in, minus, the dropout rate, graduation rate and other variables. Past the initial enrollment process, student populations are never static, they are in a constant state of flux.

1.1.2. Macro-Educational Models

The second category of educational models, macro-educational models, are those examining the educational system as a whole external to the classroom situation and experiences. Here, emphasis is placed upon total quantities whether they are incoming number of students, required classrooms, teachers, etc. They are concerned with the quantitative levels of admission or enrollment to an educational system through which a student can pass. Many important considerations of the planning process such as school curriculum, quality of instruction, etc., are omitted in these models. These models help us explain initial enrollment but contribute little to understanding issues like attrition, retention and graduation rates. The work by Bruggink, and Gambhir, in 1996 is a classic example of the application of a macro-educational approach [10].

1.1.3. Hybrid-Educational Models.

Hybrid models are, obviously, blends of the micro and macro model types. These models are more holistic in their nature because they are willing to examine a broader range of both quantitative and qualities variables. They are more methodologically complex than the micro or macro models and often result in mathematical equations that describe an environment and to some degree can predict variations in the environment. Some of the most complex of these are considered "simulation" models where the researcher can input a change in one of the independent variables and it will estimate the impact on the dependent variable. Obviously, these simulation models can only be effectively developed in environments where large quality data sets like freshmen attrition are readily available. The other primary criticism of the simulation models, is that like most modeling based on multivariate equations, it adds little to our human perception of what is occurring. An excellent example of a hybrid model is presented in *A predictive model of inquiry to enrollment* [11]. These models may be able to effectively predict but can rarely be translated into a cognitive theory that is interpretable.

1.2. Mathematical Based Prediction Models

In addition to conceptual model considerations, it is also important to examine the possible mathematical approaches that could be used in population modeling. There are a great diversity of mathematical approaches that can be used for population projections. The range extends from models based on simple arithmetic manipulations, to models requiring very sophisticated multivariate mathematical computations in order to obtain solutions. Within this range, three distinct mathematical approaches can be identified. They will be referred to here as, Deterministic models, Markov Chain models and Mathematical Regression models. Each of these types will be briefly described however, no interpretation should be made of the order in regard to sophistication or merit. All three approaches have yielded valuable insight, and it would be wrong to regard the list as forming a hierarchy.

1.2.1. Deterministic Models

Deterministic models do not have a sophisticated statistical basis and they are limited in the extent to which variations in a particular variable can be taken into

account. Generally, they indicate some constant relationship between two or more quantities or at least allow these quantities to grow at some constant and predetermined rate. Many deterministic models are based either on simple difference equations or on differential equation sets of varying complexity [12] (Johnstone, 1974). They do not include any allowance for specifying underlying probabilities, assumptions, or distributions. Consider, the previous example of factors that determine the enrollment of first grade in a given school. A deterministic model could be constructed for this specific prediction. For example, (aEt = aR•aPt) would be the expression for predicting enrollment in the first grade of school. It is very simplistic and assumes that all children of the appropriate age would be enrolling in public school. This equation expressed in words says that the enrollment (E) of pupils of age (a) in year (t) is equal to the enrollment ratio (R) of pupils of age (a) multiplied by the total number of people in the population (P) who are aged (a) in years. The user of the above equation must set the values for the enrollment ratio or any other variables to be included, such as graduation, transfers, and grade repetition rates before the model can operate. They can do this either by assigning them constant values or by allowing them to vary according to a predetermined function. Either way, the assumption is made that there will be some constant factor in the determination of future values of these parameters. This assumption is made without regard to mathematical probability considerations. This basic principle allows the researcher to build an equation for any specific need or environment incorporating as many variables as needed but limits generalizability since it is not based on a specific mathematical theory. Two early examples of deterministic equations used in educational projection are the 1968 work of Sisson [13] and Pollard [14] in 1970. In both studies the efforts were directed at improving the effectiveness of resource allocation much like the primary goal of this study.

A variation in the use of deterministic mathematical models is provided by simulation models previously discussed. Here, many equations, mostly of the deterministic type, are constructed to describe a specific situation. The result should be a duplication of the essence of the entire system under consideration. During this operation, ranges of values are substituted for the various parameters in the equation rather than the single, most likely, values as would be done with the simple deterministic model. Thus, a simulation model differs from the other model types discussed in this paper in that it is not primarily concerned with predicting final numbers but with assessing the effects and implications of various decisions on the functioning of a system.

1.2.2. Markov Chain Models

Markov Chain models are those developed on the statistical theory of Markov Chains. Markov Chain theory is based on a unique variation of basic probability theorems [15,16]. It must possess a property that is usually characterized as "memorylessness": the probability distribution of the next state/step depends only on the current state and not on the sequence of events that preceded it. This specific kind of "memorylessness" is called the Markov property. The states between which

transition is possible, could include all educational levels, as well as states such as the world outside of the educational system. Hence, not only are promotion and repetition rates included in the matrix but also dropout, graduation, death rates, population expansion rates, etc. Deterministic models can produce the same result as Markov Chain models. But, since the latter can be manipulated in accordance with the extensive theory developed explicitly for Markov Chains, these models might be considered to have better generalizability even if they are much harder to develop. There is little evidence of Markov Chain equations being used for effective prediction at the institutional level. The theory is designed for much larger population than are encountered at the institutional level and most attempts with smaller populations have met with limited success [17].

1.2.3. Mathematical Regression Models

The remaining approach, regression models, is clearly the most statistically valid model and the type familiar to most researchers. Regardless of the specific regression equation, the approach is similar. A large historical data set is analyzed for patterns and the patterns are used to project into the future. Unfortunately, as previously discussed, the historical data sets to support even a basic regression model are not available for the study area in question. While it would be possible to utilize a regression equation from a preexisting similar population any statistical validity would be lost. Thus, some form of deterministic equation is the most viable prediction approach.

1.3. Existing Literature on Prediction Models

The values of some of the independent variables required for the differential equation will have to be obtained from the existing literature. Since attrition rates and other similar variables cannot be calculated for a population that does not exist, values for these will "estimated" from previous research. It is recognized that such generalization is fraught with peril, but as previous researchers have observed, you sometimes need study questions that are not methodology clean. All population projections are to some extent a leap of faith no matter how many quality data sets you have available.

Many educational prediction models in the literature would best be described as hybrids of the micro and macro approaches. Dursun demonstrated in her 2012 research that given sufficient data and the proper variables, data mining methods are capable of predicting freshmen student attrition with roughly 80% accuracy [18]. Her research utilized a popular data mining methodology called CRISP-DM (Cross Industry Standard Process for Data Mining). This model utilized both micro data like class subject matter or class enrollment patterns and macro data like admission scores and previous grade point average in high school. One of the major contributors to the prediction accuracy of her method was the amount of available data on freshmen attrition. Attempting such a methodology with the much sparser data sets related to doctorial attrition, would be very unlikely to produce such accuracy.

Because admission patterns and successful degree completion rates at the graduate level vary significantly

according to the field of study, the most beneficial previous research will come from the discipline of criminal justice/criminology or similar social science disciplines [19]. Lightfoot and Doerner in 2008 examined completion rates, risk factors, and successful strategies of a single graduate school's programs in criminal justice and criminology [20]. Their research examined admission standards, student characteristics, and institutional traits in an attempt to understand attrition. Their findings showed that subjects with lower GRE scores, students who are non-White, and those individuals who did their master's studies at the same institution are more likely to complete the Ph.D. Overall, the doctorial attrition rate in their study was a little over one third of all admissions. Their results are in general agreement with the early data on criminal justice doctoral graduation rate of 50% [21]. When this data is considered along with data from archival data bases on similar social science programs it becomes somewhat possible to estimate retention, one of the important variables for a deterministic equation on the population in question.

Another one of the independent variables that can be partially examined in the literature is the yield rate from doctoral student applicants. In 1990, the Journal of Criminal Justice Education published the first know data on the application to yield rates for 34 existing criminal justice doctoral programs. While the data varies significantly between institutions, this early data set lets us examine overall yield rates and the relationship between perceived prestige and yield rates. The data shows the yield rate running from approximately 20% up to 88% with an average of about 40%. This data must be approached with caution since there is no assurance that the institutions gathered the data in exactly the same manner. Application yield rate is another important variable for the construction of the deterministic equation.

In addition to showing some general relationship between program prestige and yield rate, another pattern is noticeable in the 1990 data. Geographically remote programs have a lower yield rate than more centrally located programs. While not specifically studying criminal justice, this pattern was empirically verified by the research of Chen et al in their study of the spatial enrollment pattern of a pharmaceutical program [22]. Chen and his colleagues demonstrated a strong geospatial relationship between residence location and program (83% of enrollment came from within 200 miles of the program). Again, when combined with existing resident vs non-resident data from similar social science programs, a strong argument can be made for geography as an independent variable or at least a variable that would be influencing the "applicant yield rate" variable.

2. Methodology

For the prediction model in this study, many of the independent variables values are "estimated" from the existing literature as previously discussed. This is clearly less accurate than the more established method of analyzing historical data to identify the independent variables. Since there is obviously no historical data on a new proposed program to analyze, it is the best of the bad options. While there are assuredly many independent

variables that were never identified or quantified the process of predicting enrollment is at its foundation dependent on knowing the nature of the population that will be recruited from as well as identifying the variables that impact the recruitment process.

2.1. Determining the Recruitment Base

Doctoral program applicant yield rates, retention rates and the other variables that impact currently enrolled population predictions are of little utility if you can't reasonably estimate the number of possible applicants. Currently, the literature gives us only historical data on what impacts the admission process and does little to address the "applicant pool" or the question of how many qualified people are interested in entering a new doctoral program in criminal justice. The Association of Doctoral Programs in Criminology and Criminal Justice surveyed existing doctoral programs in the nation during the spring of 2000 and found that there were 627 applications to doctoral programs in 2000. However, there is no way of knowing the number of actual individual applicants. Does this number represent 627 unique applicants or a 100 applicants applying to six programs each? It also tells us nothing about the number of interested applicants who did not apply because of limited access to programs.

As is the tradition in the social sciences, for this study we will attempt to address this issue by surveying possible recruitment populations with a variety of techniques. The groups are not all inclusive, instead they represent a known set of individuals that have the general characteristics that would make them possible successful applicants to a new doctoral program upon which data might be reasonably obtained. It is also recognized, that there is overlap in the populations and survey questions were used in an attempt to identify the overlap so it could be controlled for. The following groups were surveyed for this study:

1. National Alpha Phi Sigma (APS) population
2. State (Texas) Alpha Phi Sigma population
3. Current Texas State high GPA (3.25/4.0) Alpha Phi Sigma Undergraduate BSCJ Students
4. Current Texas State high GPA Non-Alpha Phi Sigma Undergraduate BSCJ Students
5. Current Texas State MSCJ Students
6. Texas State MSCJ Student Alumni from the past 5 years
7. Texas State high GPA BSCJ Alumni from the past 5 years
8. Texas State high GPA General Social Science Alumni from the past 5 years
9. Local (75 mile radius) Criminal Justice Professionals with advanced degrees

The surveys were traditional mail based with the exception of the survey of existing criminal justice professionals. The state and national APS survey had to be conducted by mailing the surveys to the chapter offices since their policy at the time (2007) prohibited releasing individual membership information. The other student populations were traditional individual mailings based on institutional records that were believed to be accurate. The response rate ranged from almost 25% on the national APS survey to over 50% on the survey of current MSCJ students and recent MSCJ alumni. It was recognized that

this is not an exhaustive list of possible sources of student enrollment, but it did not seem reasonable to expect that other schools would provide us with a list of the names and addresses of recent high GPA graduates.

The working professional's survey was web based utilizing a snowball sampling technique. This technique was chosen because there was only a very limited sample (N=44) of known working professionals with advanced degrees available with electronic mailing addresses. Known professionals were sent the web survey with the instructions to forward the survey to other professionals they believed to be qualified within the specified 75 mile radius. The server was set to collect IP address and routing data on the surveys to indicate if we were receiving survey data from outside the specified geographical area or from the same respondent more than once. After the survey period was closed the IP addresses and routing information files were deleted to protect the anonymity of the respondents. For obvious reasons, no response rate can be calculated on the web survey. It should be noted that while the web survey was anonymous, responses from duplicate IP addresses and routing data from outside the region caused the deleted of 6 subject responses. The written and electronic survey instruments collected information on the applicant's qualifications, interest in doctoral education, doctoral programs of choice, considerations in program selection, financial, geographical, etc., and their current state of educational planning.

2.2. Building the Differential Equation

To project the enrollment of a new Criminal Justice Ph.D. program for a period of time involves a number of variables that must be included in the deterministic equation. Some of these variables can be reasonably well quantified and some are very difficult to quantify. These variables impact populations in a way that might produce student enrollment or that would cause enrollment loss. This discussion attempts to accurately quantify as many of the variables as possible, thereby reducing the number of "estimates" that must be included in the projection. Even with this attempt at quantification there are a significant number of variable that that must be "estimated" through professional judgments. These "estimates" are identified and the reader can determine the reasoning behind the professional judgment. To simplify the issues, the sources of possible enrollment are discussed separately from the sources of possible student loss. In a case like the present one, there is no historical experience so it is necessary to substitute the historical experiences of other similar programs.

In summary, for each specific identified subpopulation an equation is set up which utilizes the best available data to estimate the program yield from that specific group. The population yields are summed to produce a total expected enrollment. To make the presentation more understandable an example of a single population (National Alpha Phi Sigma) is presented in Table 1 and discussed.

Population name. This row identifies each of the subpopulations that are included in the projection. A detailed description of each of the populations is available upon request.

Table 1. Single Population Example

Population Name	Alpha Phi Sigma (APS) National
Estimated Size	1376
Availability Divisor	37% X 1376 = 509
Renewal Rate	100%
Applicants/Applications	509 X 14.8% = 5/225
TX State Applications	1.8% X 225 = 4.1
TX State Acceptance	34% X 4.1 = 1.4
TX State Enrolled 54% Yield	1.4 X .54=.76
Full Time Part Time Ratio	100/0
New FTE Students Per Year	.76

2.2.1. Estimated Population Size

The estimated size of each population is created from using the best available records or survey data for the population. This Table is believed to be generally accurate for most groups. There is a possibility of some overlap in group membership. For example, a member of the Alpha Phi Sigma (APS) National Population could have also been included in one of our Alumni Populations.

2.2.2. Availability Divisor

Not all of a surveyed population is actually ready to apply to a doctoral program at the time they are surveyed. For example, the APS population contains subjects that range from second semester sophomores to subjects in their last semester of a master's program (most were seniors or graduate students). A divisor factor of 37% was used, based on the survey responses, to estimate the number that would actually be available to apply in any one year. Different populations will have different availability divisor factors. Some historical populations, like local professionals, would have a 100% availability since they were required to have a master's degree to be in the surveyed population.

2.2.3. Renewal Rate

Some populations are renewable at a very high rate and some are renewable at a lesser rate or non-renewable. For example, the APS population is considered 100% renewable. This means that as 37% of the subjects "exit" the population they are replaced with subjects coming in as new members. Other populations, like local working professionals, may only have a 5% renewal rate. Highly renewable populations are the most valuable since they represent a sustained source of potential applicants while low renewable rate populations can only be relied upon for initial population needs.

2.2.4. Applicant/Applications

Applicants were determined by multiplying the percentage of survey respondents indicating they would like to pursue a Ph.D. in criminal justice by the estimated population. Applications are simply the number of applicants multiplied by three. This is based upon the assumption that the average applicant will apply to three schools. On the surveys, the subjects were offered the option of selecting up to five possible schools to apply to. However, the average respondent selected slightly less than three schools on the survey.

2.2.5. Texas State Applications.

This row represents the market share (percent) for Texas State as determined by the survey data multiplied by the number of applications. In the APS National example 1.8% of all the selections indicated an interest in Texas State, this percentage is multiplied by the number of possible applications. This percentage share is comparable to Sam Houston State University (a similar university) which received a 1.7% share with a 30 year history of operating a doctoral program. Market share also changes with surveyed populations. For example, the Texas State market share is significantly higher when considering only Texas APS members. This is why the APS populations were divided into Texas and National (non-Texas) for the projections equations.

2.2.6. Texas State Acceptance

The Texas State acceptance row is based upon the assumption that the Texas State program will accept 34% of the applicants that it receives. While there are no significant data sets on acceptance rates for criminal justice doctoral programs, rates are obtainable for a composite of social science doctoral programs. A data set was obtained from the University of California, San Diego (UCSD) on their 2002 applications for all Social Science doctoral applicants N=1449. UCSD is a campus of approximately 24,000 students which is significantly smaller than Texas State at 36,000 but which has a greater emphasis on graduation education than Texas State. The program at USCD, which is the most closely related program to criminal justice (sociology) has an acceptance rate of 24%. In addition, a phone survey of six similar criminal justice Ph.D. programs was conducted to determine their acceptance rates. In this survey the comparable institutions had a 34% acceptance rate. Therefore, the adoption of a 34% acceptance rate for our program seems reasonable for traditional graduate students

2.2.7. Texas State Enrolled

Traditionally only a portion of the students accepted to a program actually enroll in the program. According to the USCD data for all social sciences 34% of those students accepted actually enrolled. The nature of an enrolled population is the interaction between those accepted and those who actually enroll. High prestige programs might produce a higher actual enrollment rate and lower prestige programs will, logically, produce a lower enrollment rate. In the national survey data, Texas State was the 17th most frequently chosen program of the 27 programs that could be selected (survey data available on request). This would indicate that Texas State University has a middle prestige position. In a survey of the acceptance rate of other similar criminal justice Ph.D. programs it was discovered that the average enrollment rate was 54%. Since Texas State had achieved a middle prestige ranking without any promotion the use of a 54% enrollment seems reasonable.

2.2.8. Full Time/Part Time Ratio.

Based upon historical experiences, the various populations have been classified as to their respective percentage of part time vs. full time students. Traditional populations like the APS populations are considered to be completely full time (often assistantship funded). Regional populations, which are more likely already employed, are considered to be 75/25% full time to part time. Local working professionals are considered 25/75% full time to part time ratio.

2.2.9. New FTE Students Per. Year

As the last row in Table 1 indicates, when the math is done, the national APS population should generate only about ¾ of a full time student (18 semester credit hours per year). Table 2 shows the calculations for all of the various populations as a single table. This calculation would be for the first actual year population only. Since the populations variables will have different values in different years.

Table 2. Population Source Example

Population Name	Alpha Phi Sigma National	Alpha Phi Sigma Texas	CJ High GPA Non-APS National	CJ High GPA Non-APS Texas	Repeat for each of the 9 groups*
Estimated Size	1376	156	5504	624	*
Availability Divisor	37% X 1376 = 509	37% X 156 = 58	37% X 5504 = 2036	37% X 624 = 231	*
Renewal Rate	100%	100%	100%	100%	*
Applicants/Applications	509 X 14.8% = 5/225	58 X 14.1% = 8/24	2036 X 14.8% = 301/903	231 X 14.1% = 33/99	*
TX State Applications	1.8% X 225 = 4.1	13.5% X 18 = 3.2	1.8% X 903 = 16.3	13.5% X 99 = 13.4	*
TX State Acceptance	34% X 4.1 = 1.4	34% X 3.2 = 1.1	34% X 16.3 = 5.5	34% X 13.4 = 4.6	*
TX State Enrolled 54% Yield	.76 students per year	.59 students per year	2.97 students per year	2.48 students per year	*
Full Time Part Time Ratio	100/0	100/0	100/0	75/25	*
New FTE Students Per Year	.76	.59	2.97	2.17	*
Total of All Population for First Year					20.19*

For example, your local professional market will be heavily exhausted in the first admission cycle and produce a much lower yield for each subsequent year. As the program enters its second year a similar equation is used to introduce attrition into the population. Attrition is comprised of academic program failure, voluntary drop-out rate and graduation rate. These have to be calculated for each program year since the variables value will change by year. There would obviously be no graduation in the first year but some graduation should start to appear in the third or fourth year. The combination of projected enrollment data and attrition data is used to produce Table 3. Presented in this Table is the projected FTE enrollment for

the first 10 years with the actual FTE enrollment for the five years of time the program has been in operation.

Table 3. Projected VS Actual Ph.D. Enrollment

	New FTE Projected	Projected Attrition FTE	Projected Graduates FTE	Total Projected FTE	Total Actual FTE
Year 1	20	0	0	20	13.3
Year 2	14	2	0	32	20.2
Year 3	9	3	1	37	34.2
Year 4	9	3	4	39	38.9
Year 5	9	4	4	40	41.2
Year 6	9	4	5	40	N/A
Year 7	9	4	5	40	N/A
Year 8	9	4	5	40	N/A
Year 9	9	4	5	40	N/A
Year 10	9	4	5	40	N/A

In comparing projected to actual enrollment the reader will note that the actual enrollment was significantly lower than the projected enrollment for the first two years of operation. There are several possible explanations for this difference, including the obvious one that the projection was faulty. In planning for the program, it was anticipated that there would be time to advertise, select and offer funding to a number of out of state and international applicants for the first year. However, due to decisions made outside of the academic unit, the program was started up on short notice (approximately 45 days) preventing a traditional academic recruiting and admission cycle. The first student cohort was primarily made up of working professionals and students available in the geographic area. The short lead time would not have accommodated out of state or international students relocating.

Following the data projection example used in Table 2, the out of state and international contribution to the population should have been an additional 5.99 FTE. If that projection is added in to the actual FTE enrollment of 13.33 it would have produced an enrollment almost exactly on the projection of 20 FTE. This hastily recruited cohort also had a greater than expected attrition rate, especially among the working professionals, that would continue to impact the projection into the second year. Being a full time working professional and significantly participating in a doctoral program turned out to be less practical than many of the early students had assumed. In the subsequent three years, the FTE projection was remarkably close to the actual enrollment. It should be noted that the author played no significant role in the admission decisions during this five year period. In addition, these projections were submitted to an external state agency and presented at the 2007 Academy of Criminal Justice Sciences annual meeting in Seattle (Panel 255), several years before any admissions decisions were made.

3. Conclusions and Recommendations

In conclusion, the author believes that this project successfully demonstrates the effectiveness of combining survey data and archival data with a deterministic equation to produce reasonably accurate projections of new doctoral program enrollment. As to suggestions for future research, the author believes that there is great potential in the application of Markov Chain theory to this type of population projection. To effectively implement a Markov Chain equation would require a much larger population size and an increased number of independent variables. While no single institution could probably achieve the required population size, a data set of a number of similar institutions combined might support the equation and produce a more sophisticated projection model. The development of a Markov Chain equation was beyond the scope of the current project. The current project, first and foremost was designed to satisfy the administrative needs of the university in the decision to fund a new program.

Acknowledgement

References

[1] Dickeson, R. C. (2010). *Prioritizing Academic Programs and Services* 2nd Ed. Jossey-Bass, Inc./John Wiley & Sons, Inc.

[2] Clagett, C. (1992). "Enrollment Management." In M. A. Whiteley, J. D. Porter, and R. H. Fenske (eds.), *The Primer for Institutional Research.* Tallahassee, Fla.: Association for Institutional Research.

[3] Brinkman, P. T., & McIntyre, C. (1997). Methods and techniques of enrollment forecasting. In D.T. Layzell (Ed.) *Forecasting and managing enrollment and revenue: an overview of current trends, issues, and methods* (pp. 67-80)..

[4] Lightfoot R. C., & Doerner W. G. (2008). "Student Success and Failure in a Graduate Criminology/Criminal Justice Program" *American Journal of Criminal Justice*. 33(1):113-129.

[5] Weismann, J. (1994). "Enrollment Projections: Combining Statistics and Gut Feelings." *Journal of Applied Research in the Community College*, 1994, 1 (2), 143-152.

[6] Correa, H. (1967). A survey of mathematical models in educational planning. In *Mathematical models in educational planning*. http://files.eric.ed.gov/fulltext/ED024138.pdf#page=22.

[7] Bush, R. R. & Mosteller, F. (1955). *Stochastic Models of Learning*. John Wiley & Son, New York, NY.

[8] Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, 77,481-495.

[9] Scandura, J. M. (1970). Development and evaluations of individualized materials for critical thinking based on logical inference. Rending Research. *Acta Psychologica*, 63, 301-345.

[10] Bruggink, T. H., and Gambhir, V. (1996). Statistical models for college admission and enrollment: A case study for a selective liberal arts college. *Research in Higher Education* 37(2): 221-240.

[11] Goenner, C.F., Pauls, K. (2006), "A predictive model of inquiry to enrollment", *Research in Higher Education*, Vol. 47, No. 8, pp. 935-956.

[12] Johnstone, J. N. (1974). Mathematical Models Developed for Use in Educational Planning, *Review of Educational Research* Spring pp. 177-201 American Educational Research Association.

[13] Sisson, R. L. (1968). A hypothetical model of a school. Pennsylvania University, *Fels Institute of Local and State Government*. ERIC No. ED030978.

[14] Pollard, A. H. (1970). Some Hypothetical models in systems of education. *The Australian Journal of Statistics*, 12, 79-81.

[15] Kemeny, J. G., & Snell, J. L. *Finite Markov chains*. Princeton, New Jersey: D. Van Nostrand Co. Inc., 1960.

[16] Kemeny, J. G., & Snell, J. L. *Mathematical models in the social sciences*. Boston, MA: Ginn Publishing Inc., 1962.

[17] Harden, W. R. & Tcheng, M. T. (1971). Projection of enrollment distribution with enrollmentceilings by Markov processes. *Socio-Economic Planning Sciences*. 5, 467-473.

[18] Dursun, D. (2012). Predicting Student Attrition with Data Mining Methods. *Journal of College Student Retention*, Vol. 13(1) 17-35, 2011-2012 Baywood Publishing Co., Inc.

[19] Bowen, W. G., & Rudenstine, N. L. (1992). *In pursuit of the Ph.D.* Princeton, NJ: Princeton University Press.

[20] Lightfoot R. C., & Doerner W. G. (2008). "Student Success and Failure in a Graduate Criminology/Criminal Justice Program" *American Journal of Criminal Justice*. 33(1):113-129.

[21] Klyman, F. I. and Karman, T. A. (1974) A Perspective for Graduate-Level Education.*Criminal Justice Crime & Delinquency* 1974 20: 398.

[22] Chen, K., Kennedy J. Kovacs, J. M. and Zhang C. (2007). A spatial perspective for predicting enrollm.ent in a regional pharmacy school. *GeoJournal* 70:133-143.