DEPARTMENT OF OFFENDER REHABILITATION

LOUIE L. WAINWRIGHT, SECRETARY

Research Study

COMPARATIVE ANALYSIS OF POPULATION PROJECTION METHODS

Document #77-R-064

Bureau of Planning, Research and Statistics

September 27, 1977

67412

FLORIDA DEPARTMENT OF OFFENDER REHABILITATION
LOUIE L. WAINWRIGHT, SECRETARY

COMPARATIVE ANALYSIS OF POPULATION PROJECTION METHODS

SEPTEMBER 27, 1977

Robert Roesch, Bureau Chief
Lonnie Fouty, Administrator of Research
Robert Morgen, Intern

TABLE OF CONTENTS

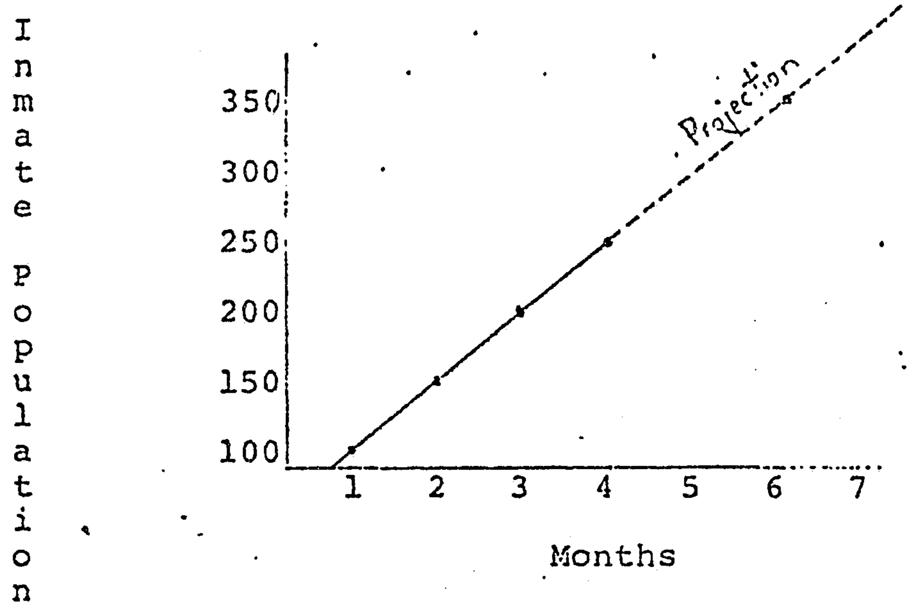
PERCENTAGE OF INCREASE AND LINEAR REGRESSION

Effective and efficient correctional planning depends to a great extent on the ability to accurately predict the future inmate population. In response to this need the Department of Offender Rehabilitation has spent considerable time in attempting to develop more satisfactory projection techniques.

For several years prior to 1973, the Department used the simple technique of projecting the past percentage of change to estimate future population. The assumption under this method was that the population would change by the same percentage in the future as it had done in the recent past. Thus, if the inmate population had grown by 4% over the past year, the projected growth rate for the future would also be 4%. In this technique, only two data points are used in making projections.

In 1973, the Department began to utilize the technique of linear regression, because of its promise of greater accuracy. In the application of this technique by the Department, the past values for the inmate population (the total number of inmates at the end of each month) are used to predict the future values. Graph 1 illustrates the operation of this technique.

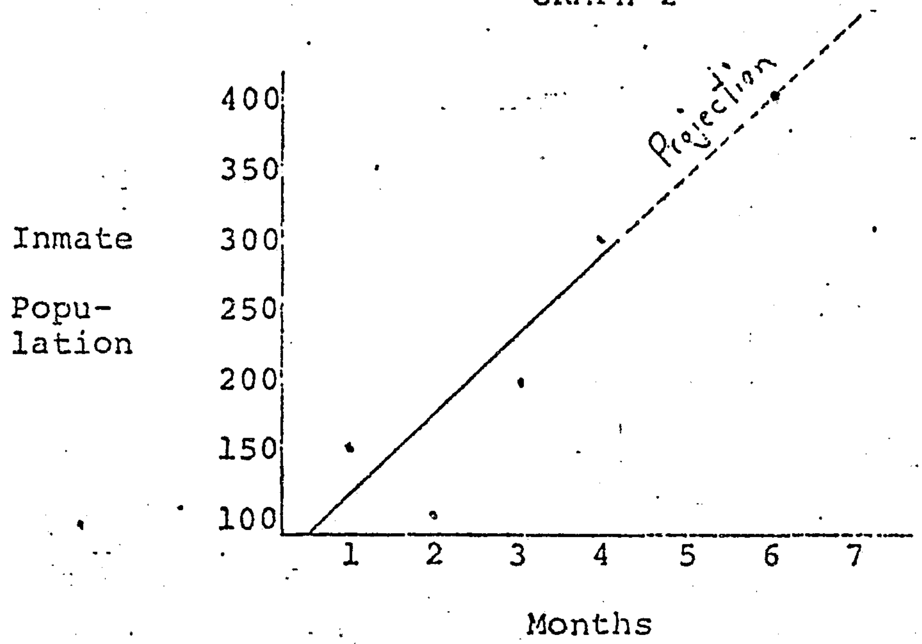
GRAPH 1



The graph shows that the inmate population increased by 50 each month (solid line). There were 100 inmates the first month, 150 the second month, 200 the third month, and 250 the fourth month. Linear regression takes the known past values and projects these into the future. Since the inmate population has grown by 50 each mohth, it is projected that the population will continue to grow by 50 each month. Thus the population for the sixth month, for example, would be expected to reach 350.

The past data does not have to form a straight line, as in Graph 1. Graph 2 illustrates a possible array of past data points:

GRAPH 2



The object is to make the distance to the points above and below the line as small as possible. The projection of the future population is done in the same way as in Graph 1, that is, by extending the line forward.

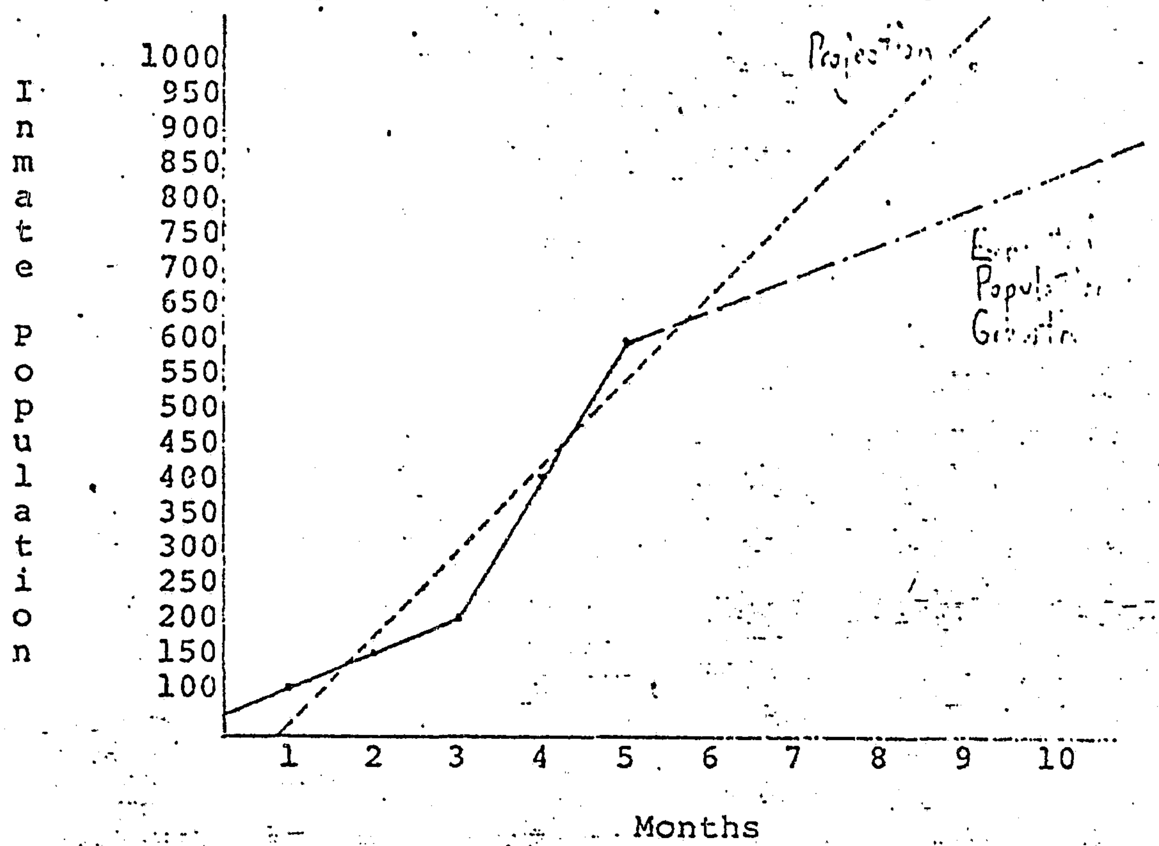
Thus, in linear regression the prediction of future events is wholly based on a series of historical events. One of the advantages to this technique is the relative ease of data collection required to generate projections. In the case of inmate population, all that is needed are the past figures on the population that are readily available to the Department of Offender Rehabilitation.

Another advantage of linear regression is that it can provide reasonably accurate predictions of the future inmate population so long as a particular trend has been stable in the past, and is expected to continue. What the trend is does not affect the accuracy of the technique, so long as it is a stable and linear one. For example, there may be a continuous steady increase in the inmate population, or a continuous gradual decline. The primary assumption of the linear regression technique is that whatever has happened in the past will continue to happen in the future.

Obviously, real systems do not necessarily perform in linear fashion. There are times when the inmate population fluctuates widely just as there are times when a stable pattern shifts abruptly. Under these conditions linear regression is much less useful and may in fact be highly misleading. This is the results of limitations of the linear regression technique in responding to major changes in critical variables.

Graph 3 illustrates a situation in which linear regression will not produce an accurate estimate. In this case there has been a slow pattern of growth for the first three months, followed by a sudden shift to a very rapid rate of growth.

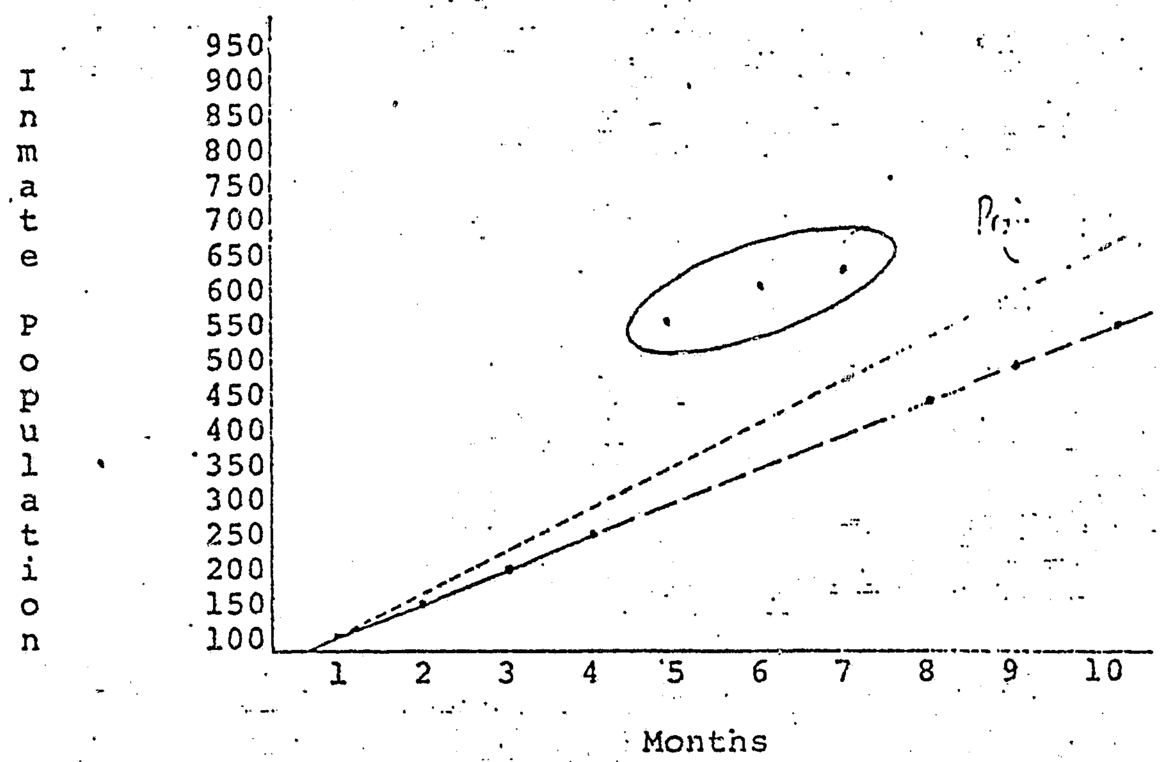
GRAPH 3



If the data points for the first five months are used, the projection shown above results. This is because the line is always drawn so that the distance of the points above and below it are minimal. In the above case, the use of the linear regression technique causes an overestimation of the future inmate population. As can be seen in Graph 3, the inmate population returned to a slower rate of growth after the fifth month.

Linear regression is also sensitive to outlyers. Outlyers are extreme values in a array of points. Graph 4 shows this case:

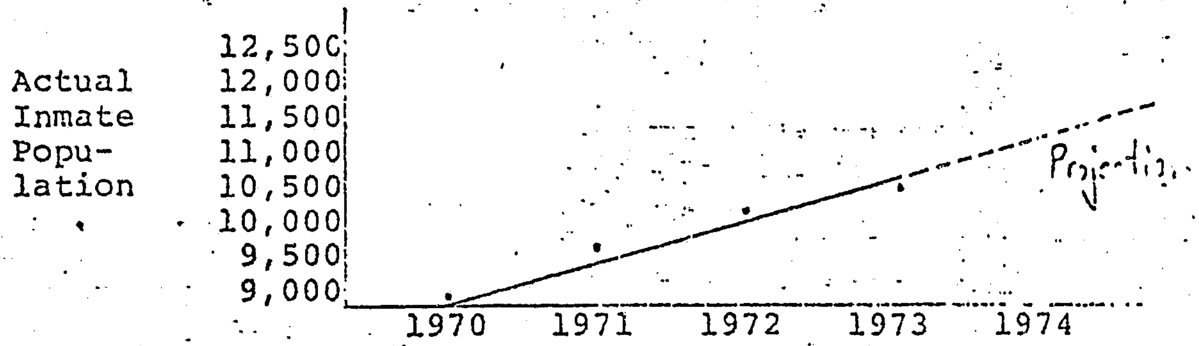
GRAPH 4



The points circled represent the outlyers. The array of points shows constant growth for the first three months, rapid growth for the following two months, and a return to constant growth. The regression line will fit the array of points so that the distance of the points above and below the line is minimal. It is clear in Graph 4 that the projections will be higher than the actual inmate population. The reason is that the extreme values (outlyers) pulled the projection line up. The dotted line represents the expected growth of the inmate population.

How well has linear regression performed in the recent estimates of inmate population? This has depended on whether or not the rate of growth was constant. For instance, if linear regression had been used to predict the inmate population prior to 1974, the estimates would have been very accurate because the rate of growth was constant.

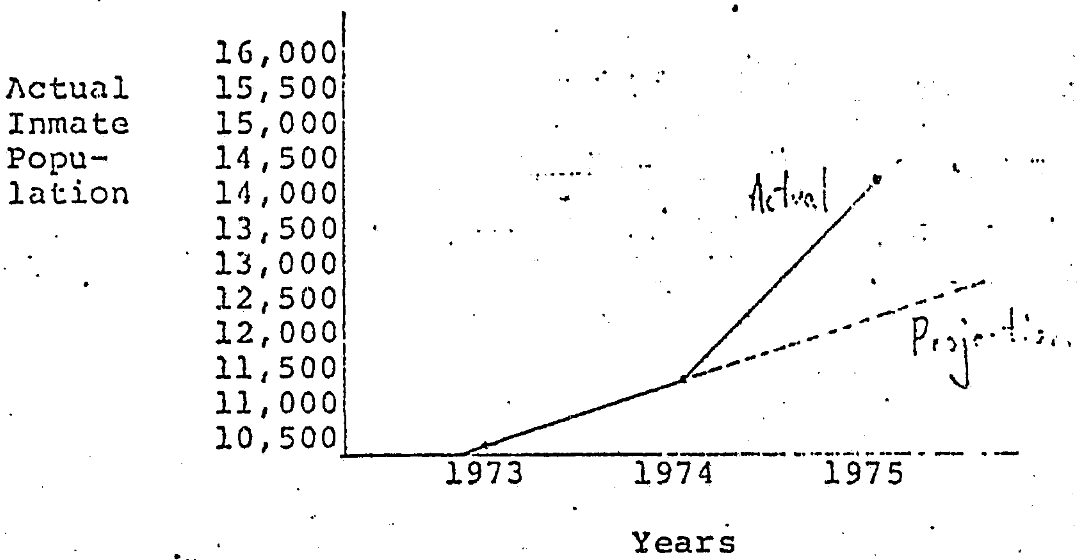
GRAPH 5



Graph 5 shows a constant increase in inmates from 1970 through 1973.* Projections of the 1974 inmate population approximated the 1974 actual inmate population because the 1974 inmate population increased at the same rate as did the 1971, 1972, and 1973 inmate population. In this case, linear regression was very accurate.

During 1975 the inmate population increased dramatically. Projection of the 1975 inmate population, using data bases prior to 1975, did not approximate the actual 1975 inmate population, as can be seen in Graph 6.

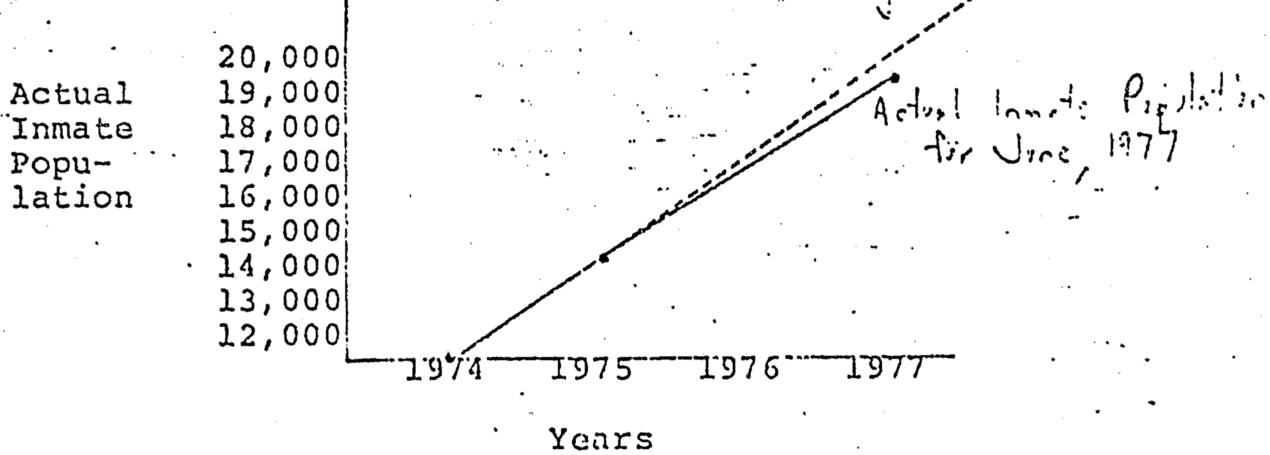
*In its use of linear regression, the Department actually used a large number of data points, each one of which represented an end-of-month population. The presentation here of only a few points representing yearly population is made only for purposes of illustrating the technique.



Linear regression estimated the inmate population for June, 1975 to be approximately 12,000. The actual inmate population for June of 1975 was 16,066. In essence, the projection underestimated the 1975 inmate population because the data base showed slow growth (thus predicting slow future growth), whereas the inmate population actually increased rapidly.

By contrast, linear regression using end-of-month population in 1975 as a data base overestimated the 1976 inmate population (Graph 7).

GRAPH 7



The method projected the inmate population for June, 1977 to be 20,000. The actual inmate population at that time was 19,534.

Similarly, if the rate of growth of the prison population begins to decrease, projections using the 1975 or 1976 inmate populations will overestimate future populations. In order to circumvent these problems, rolling data bases or continuous updates were implemented.

II

THE SIMMODG MODEL

As we have indicated, the linear regression technique uses historical events exclusively in the prediction of future events. In the case of inmate population, the historical events used are the past figures for total inmate population. The future inmate population is seen as following directly from the past inmate population. Fairly representative predictions are possible when the historical trend has been stable, and linear in nature, but not when sudden shifts occur.

The problem with using the linear regression technique for the prediction of the future inmate population is that there are many volatile factors which may cause major shifts in the population level. The crime rate, for example, may have held steady for the period of time covered by the projection, thereby increasing user confidence in the outcome. This does not mean, however, that these conditions will continue unchanged over an indefinite period. On the contrary, the crime rate may suddenly start increasing at a rapid pace. Using linear regression, there is no way to anticipate these important changes.

Faced with the obvious limitations of the linear regression technique, the Department of Offender Rehabilitation surveyed and utilized other techniques in an attempt to improve the confidence level of predictions developed by the Department.

The SIMMODG model was developed in early 1976. Rather than simply examining the past inmate population, the SIMMODG model considers the criminal justice system as a whole. This model examines the interrelationship of prison, parole, probation, and arrests.

Although the model does not allow long term predictions, it does make better short term predictions possible. The basic model is formulated as a set of linear estimations predicting prison, parole, and probation populations at time (i + 1) using rates and population levels at time i.

$$N_i + 1 = \lambda A_i + rP_i + vL_i + (1 - \mu) Ni$$

$$P_i + 1 = bN_i + (1-r-q)^{P_i}$$

$$L_i + 1 = \sigma + (1-v-w)^{L_i}$$

The variables N_i P_i, and L_i represent the populations for prison, parole, and probation, respectively; A_i denotes the number of arrests at $time_i$. The time increment is usually chosen to represent one month. Besides the initial values of N_i, P_i, and L_i at time i = o, the remaining constants are calculated using the historical data available in the model data base.

$$\lambda = \frac{\text{new prison admissions}}{\text{total arrests}}$$

$$r = \frac{\text{parole violator admissions}}{\text{number on parole}}$$

$$\mu = \frac{\text{total released}}{\text{prison population}}$$

$$v = \frac{\text{probation violator admissions}}{\text{number on probation}}$$

$$c = \frac{\text{released to parole}}{\text{total released}}$$

$$\gamma = \frac{\text{released from parole}}{\text{number on parole}}$$

$$\sigma = \text{new additions to probation caseload}$$

$$w = \frac{\text{released from probation}}{\text{number on probation}}$$

$$b = \frac{\text{released to parole}}{\text{prison population}}$$

Means for the above ratios are computed using the time period indicated while running the model. For the constant λ, admissions per arrests, a lag time is associated between arrests and admissions; lag time is defined to be the expected time between arrest and incarceration. It is also possible to add another key variable, crimes reported to the police, to this simulation model.

The data needed for the use of this model is relatively easy to obtain. Data on crimes and arrests are taken from the Uniform Crime Reports, while data on the parole, probation, and prison systems are found in the Monthly Management Report and Monthly Field Recap Report, published by the Department.

For the SIMMODG model, the projections are valid for only 6 months, since the primary variable is the arrest rate/ time lag (which at this time is estimated to be six months between arrest and commitment). Any projections exceeding 6 months are based on estimated arrest figures, and are thus less reliable.

The main advantage of the SIMMODG simulation model over the linear regression technique is that important changes in the prison population can be anticipated to some extent. The linear regression technique discussed earlier in this paper does not enable the Department to foresee shifts in the prison population because it assumes that past trends will continue.

With the SIMMODG model, on the other hand, the impact of changes in the criminal justice system (especially arrest rates) on the prison population can be gauged.

This model represented an important advancement over linear regression insofar as it used developments in the criminal justice system instead of merely relying on the end-of-month prison population figures. Applications, however, are limited in that projections cannot be made with confidence for an extended time frame. When attempts were made to apply the model, it became apparent to the Department that projections for an extended period (three years, for example) would depend on making estimates of future arrest rates. This is, at present, very difficult to do with any degree of confidence.

III

THE SPACE MODEL

Another model that was developed for the prediction of the inmate population is the SPACE model. This model was developed by the Council of State Governments in the early part of 1977. Like the SIMMODG model, SPACE attempts to simulate the workings of the overall criminal justice system. The purpose is to trace the flow of offenders from the time of their arrest to the time that they are released from prison.

The principal purpose of the SPACE model is to predict short-range changes in the correctional population, short-range denoting a period of not more than a year and a half. The SPACE model can also be used to simulate the probable impact of policy changes (such as longer sentences, greater use of incarceration, etc).

There are some significant differences between the SPACE model and the SIMMODG model. The SPACE model uses the proportion of arrests placed on probation rather than using a constant number of new probations. Also, this model has a time lag built into each item that uses a rate, a feature that should increase the accuracy of the predictions.

This model enables short-range predictions to be made that turn out to be reasonably accurate. It is remarked that:

> The Florida test results confirmed the potential utility of SPACE as a valid population projection device. Using data for the period form July, 1972 until January, 1976, the program projected prison population forward to October, 1976, with an average margin of error of three percent.

As in the case of the SIMMODG model, the SPACE model possessed important advantages over linear regression. Once again, however, the model's usefulness was greatly limited by the fact that only short-range predictions were possible. Predictions for a period of more than a year and a half are necessary for purposes of capital improvements and budget preparation.

As a result, the SIMMODG and SPACE models were not utilized. In 1976 the Department decided to continue using linear regression with a moving data base while considerable effort was made to develop improved methodologies. In the same year, population projections were produced by applying log transformations to the data base, and also by using a quadratic equation with a sine curve. The projections that resulted from the use of these techniques indicated a huge increase in population which appeared to the Department to be unrealistic. As a result, these techniques were not given further consideration. In late 1976, work was begun on the SIMMS model. Work on this complex model is still in progress. In 1977, the Department queried the other states on their projection methodologies. It also developed and began utilizing a promising new model.

IV

SURVEY OF METHODOLOGY IN OTHER STATES

In its efforts to develop more satisfactory projections of the inmate population, the Florida Department of Offender Rehabilitation conducted a survey of population projection methodologies and approaches in all of the other states and the District of Columbia.

Our survey shows a picture as varied as one might expect from such a large number of diverse jurisdictions. Some states eschewed population projections entirely, while others had developed rather sophisticated projection models. In those states that have done projections, and that responded to our survey in some depth, there appears to be a general awareness of the considerable difficulties involved in trying to accurately predict the future inmate population.

One of the notable findings of the survey is the recognition, in many states, that the application of linear regression, using past inmate population as the sole data base, leaves a great deal to be desired. This recognition may have come about as these states experienced the same sudden and unexpected surge in inmate population as occurred in Florida. As we mentioned in Section 1 of this paper, linear regression is of questionable value when there is a dramatic shift in a given pattern. This shift--to much higher populations--occurred in many states between 1973 and 1975. The sudden increase in population resulted in crises in many states. These crises involved severe prison overcrowding, the use of emergency facilities to house inmates, and the effort to quickly build many new facilities. This situation could not be anticipated using the method of linear regression. This has led a number of states to attempt to develop more sophisticated methodologies.

These methodologies have often taken the form of either multiple regression or simulation models. The multiple regression models are still "linear," but they utilize more than one factor (as the name implies), and move away from past inmate population levels as a predictor of future population. The simulation models attempt to recreate the actual workings of the criminal justice system, or part of that system, in order to gauge the impact on inmate population levels.

Several states have decided against using a strict methodology. Instead, they have opted for a "multi-factor" approach that takes into account a large variety of factors, including some that are difficult to quantify. In this approach a range of estimates is frequently made, rather than a single estimate.

It is apparent that there is at this time no technique or approach that offers a "vision into the future" with a sure prospect of accurate population projections. Although much valuable information has been produced by this survey, it is difficult to assess the value of the competing techniques and approaches, for two reasons. First of all, since most of the promising techniques have been developed only recently, there has not been time to determine their usefulness and degree of accuracy. It may take several years to do so. Secondly, conditions in the different states are not the same. There are urban states and rural states; states with a high crime rate and those with low rates; states with a high rate of incarceration and states with a low rate of imprisonment, and so on. A technique that is effective in one state may not be effective in another state that has a different set of conditions.

Table 1 shows the work that the different states have done in the area of population projections.

V

REASONS FOR DEVELOPMENT OF THE NEW MODEL

The unexpected growth of the Florida inmate population since 1973 has greatly heightened the need for more effective forecasting tools. The inmate population, which stood at 10,669 in June, 1973, reached 19,534 by June, 1977, an increase of 83.1 percent in just four years.

The very extent of this increase--in sheer numbers--makes it essential that more accurate predictions be provided. The size of the Florida system dictates requirements for facility construction and capital outlay that are radically different from those in a state with a small prison system. In a system with 1,000 inmates, for example, an increase of 83% would mean the addition of 830 individuals. A similar percentage increase in Florida added nearly 9,000 inmates to the system. The requirements for additional bed spaces in the small system could be met by the construction of one or two facilities. In a system the size of Florida's, on the other hand, an increase of 83% in four years translates into overcrowded conditions, temporary emergency housing, the need for extensive construction and huge capital outlays.

Until recently, efforts made to predict future inmate population levels centered around the linear regression methodology. As we have stated earlier in this report, the specific application of linear regression used by the Department employed the moving end-of-month inmate population as the sole data base. While utilization of this technique did not enable the Department to foresee the recent dramatic increase in inmate population, it also fails to show when this increase will level off. In keeping with the logic of this method, it is assumed that a very high rate of growth would continue indefinitely. This is because the period of rapid increase was being used as the data base. It was planned that adjustments would be made to these projections when the increase slowed, allowing the department to avoid underestimation in the face of unprecedented growth.

The SIMMODG and SPACE models offered the possibility of more reliable predictions, since they monitored developments in other parts of the criminal justice system (such as arrest rates). The limitation of both SIMMODG and SPACE was that predictions could be made with confidence only for a relatively short time period.

During the period of rapid increase in the inmate population, speculation centered around the question as to when this rate of increase would level off. This speculation became more intense during the first six months of FY 1976-77, as a result of a slight downturn in DOR admissions, and of reported decreases in both crime and arrest rates.

A popular assumption was that those inmates admitted during the period of rapid increase in FY 1974-75 would be released at the same rate as they had entered the inmate population. It was expected that the average time to be served by inmates still in prison would approximate the amount of time served by inmates who had already been released from custody. It was therefore believed that if the average length of time served was 27 months or less, that over half the population admitted in FY 1974-75 would be released by the end of FY 1976-77. Monthly reports of net gains and losses were carefully scrutinized; each month that releases exceeded admissions, speculation grew that inmate population growth had, at last, leveled off.

It was erroneous to assume, however, that releases would continue to exceed admissions and that the inmate population would stabilize. This assumption resulted from the failure to recognize the significant differences in admissions and releases on an individual basis. Overall admissions and aggregated releases were dealt with as a homogeneous entity, and it was therefore believed that the average length of time served by those already released was representative of the time that has been and will be served by the entire DOR population. The residual prison population was not delineated, and its characteristics were not analyzed.

The significance of these actual characteristics of the inmate population may be understood if a hypothetical group of admissions and releases is examined in a simple illustration. If there are 100 new admissions and 50 releases in a given month, the inmate population would increase in size.

If the converse were to occur (100 releases and 50 admissions) resulting in a net monthly loss of 50 persons, there would be an immediate short-term decrease in the size of the inmate population. However, if DOR released 100 persons who were serving two-year sentences and admitted 50 inmates under twenty-five year sentences, the population in custody would ultimately grow over the long run as the short-term population replaced itself in future months. In fact, there would have to be a prolonged decline in admissions of inmates serving relatively long sentences before a significant leveling would occur in the growth of the inmate population, and this simply has not happened. It is estimated that it would take about twenty-five years for the inmate population to stabilize and reflect current moderate declines in admission without significant changes in sentencing policy by the courts.

VI

DESIGN OF SIMULATED LOSSES/ADMISSIONS MODEL: PREDICTION OF RELEASES

The design of the Simulated Losses/Admissions Model addresses certain characteristics of the DOR inmate population that have not been effectively considered in other forecasts.

Historically, the growth of the DOR inmate population has been analyzed in terms of gross admissions and releases. Linear projection and computation of net gains/losses treats each release and each admission as being statistically equal. Prior comparisons of numbers of admissions with number of releases were appropriate for providing static head counts but proved inadequate for making long-term or even short-term projections.

Previous inmate population forecasts did not account for the fact that the offender flow consists of a number of individual cases, each differing in length of sentence, offense, and other demographic and circumstantial characteristics that define the length of time that an offender will remain in custody.

Monitoring the numbers at intake and release without determining the length of time that offenders are likely to remain in the status population makes estimation of the size and/or characteristics of the future residual population impossible. Consideration of the characteristics of each offender that are significant to the amount of time he will actually serve is essential to determing the rate of release of both the residual population and the new admissions.

For example, the admission of an offender with a twenty-year sentence cannot be accurately compared (on a one-to-one, gross release subtracted from gross admission basis) to the release of a person with a three-year sentence. In terms of the size of the residual inmate population, the admitted inmate replaces the released inmate for an initial three-year period but the long-range implication over the additional time served will not be measurable under previously used methodologies.

In order to determine an indicator of the time an offender would actually remain in prison, a number of variables were examined. Among these were offense, length of sentence, race, age, and prior commitment to DOR. The highest correlation with time served (r= .66) occurred with the length of sentence.

Once it has been established that length of sentence correlates most strongly with actual time served, it is necessary to quantify the relationship so that it may be simulated. The greatest problem in case-by-case prediction of release dates is that the only data available on length of time served is data derived from the records of inmates who have already been released. This data has been significantly biased as a result of the dramatic increase in admissions over the past four years.

As a representative sample, the automated data base for releases as compared to the unprecedented number of recent admissions and the current status population is extremely limited. While there may be as many as 95% of the one-year and two-year admissions accounted for on the release tapes since FY 1974-75, the percent of releases for longer sentences is extremely small. For instance, the number of releases reported for persons sentenced to life imprisonment was 127 over the two year period. However, there are currently more than 1,650 offenders in DOR institutions serving life sentences. Of these released, the longest time served on a life sentence was about 19 years while some of those not released have served more than 30 years at this time.

In order to predict release dates, we examined: 1) the amount of time ~~that~~ has been served by those already released, 2) the amount of time that has been served by those still in prison, and 3) the estimated amount of remaining time to be served by those still in prison based upon the available sample of historical data derived from 1 and 2.

Initially, only the average length of time served by selected length-of-sentence classes was examined. The results of adding the average time served to each admission date for those in custody, counting the numbers released over time, and then comparing the release pattern over the same period with the actual releases, proved to be unacceptable. It was determined that the "standard deviation" or the amount of variance in distribution of releases from the mean (average) had to be considered. Release distribution curves were thus constructed for each of the length-of-sentence classes to be used as a basis for the simulation.

The first part of the release module was designed to produce a series of distributions of length-of-time served for fourteen length-of-sentence classes. These were used to simulate the actual rate of release of inmates and to predict future releases.

There were just over 8,000 admissions in FY 1976-77. In order to predict the number of inmates among 8,000 admissions who would be released after serving some number of months (36 months for example), the computer tapes listing offenders released during FY 1974-75 and FY 1975-76 were examined to determine the actual number released after 36 months. These inmates would have been admitted during FY 1971-72 and FY 1972-73 when admissions were about 5,000 per year. Assuming the examination of this historical data indicated that 50 inmates per year had been released after 36 months, then it would be predicted, based upon a constant proportion, that 80 out of 8,000 offenders admitted in FY 1976-77 would be released after 36 months. Since 8,000 is 1.6 times 5,000, each of these 50 inmates in the historical sample would have to be assigned a weight of 1.6 so that they would represent 80 inmates in the release distributions. Therefore, the program is designed so that a weight is assigned to each release record according to the inmate's year of admission as a means of adjusting for the inordinate increase in admissions that has occurred over the last few years.

In addition, the length of time served for each release record was multiplied by a varying factor for each length of sentence class. The reason that the length of time served had to be increased was that the unadjusted release distributions (derived from an analysis of time served by the set of inmates who have been released) were not representative of the time served by inmates still in the prison system. Especially for longer sentences, the unadjusted release distributions were based upon relatively small samples compared with the number of inmates still in prison. Some of those inmates not yet released have already served terms considerably longer than the sample of inmates who have been released and upon whose records the unadjusted release distributions are based. Adjustment of the actual release distributions was also made to compensate for unusual levels of releases in the parole sector. After the two adjustments were made for each release record, the fourteen length-of-sentence release distributions used in the simulation were generated on a monthly basis.

The next part of the program works in this manner: for the first length of sentence class, the release distribution is called N(t), where t represents the number of months to be served and N(t) represents the number of inmates who would serve exactly t months. A new function, A(t), is defined as follows:

$$A(t) = N(1) + N(2) + \dots + N(t).$$

A(t) represents the number of inmates who would serve t months or fewer. The inverse of A(t) is called T(n), where n represents a numbering of inmates to be released with respect to this distribution in the order that they would be released. T(n) represents the number of months after which the nth inmate would be released. These functions are used in the second part of the release module. Similar functions were defined for the other length of sentence classes.

The second part of the release module predicts monthly releases by assigning a predicted release date to each inmate currently in custody or admitted to the prison system. The June 30, 1973, computer status tape and the admission tapes for FY 1973-74, FY 1974-75, and FY 1975-76 were used as the data base and the program predicting the releases is called the Release Prediction Program.

For each inmate admitted to prison after June 30, 1973, the adjusted release distribution N(t), along with Λ(t) and T(n) is selected corresponding to the inmate's length of sentence. A number, K, representing the inmate to be released, is chosen at random along the vertical axis. T(K), therefore, will be the predicted length of time to be served for this inmate. Adding this time to the admission date provides a predicted release date. This type of selection assures that whenever the number of admissions is equal to the number of inmates represented by the distribution, the distribution of the predicted lengths of time to be served will be almost identical to the distribution N(t). Modifications were made for inmates already in prison on June 30, 1973, and for inmates who had received mandatory minimum sentences of three years or twenty-five years.

After some further necessary adjustments, the projected monthly populations were calculated, based on monthly admission and release figures. These were compared with the actual monthly populations. The length of sentence factors were then adjusted in order to give close monthly predictions for the four years from June 30, 1973 to June 30, 1977.

The release module is driven by projected admissions. The assumption is that the distribution of admissions in the future will be proportional to the admissions for FY 1975-76 when distributed by length of sentence and by month of admission.

VII

DESIGN OF THE MODEL: PREDICTION OF ADMISSIONS

Since the model is driven with numbers of admissions to the correctional system, it has been necessary to develop a method for predicting admissions. After consideration of many possible variables, it was found that population at risk and the state unemployment rate correlate most strongly with admissions.

Projected figures for both of these variables were readily available to the Department. Predictions on the population at risk have been made through the year 2020 by the Bureau of Economic and Business Research at the University of Florida. Projected unemployment rates through 1979 are available from the Florida Department of Administration, Economic and Tax Research Unit.

The rationale behind the use of population at risk is as follows: within the general population, there is a subset that contributes disproportionate numbers to those arrested, convicted and incarcerated. Although the exact ages to be included in this "population at risk" may vary somewhat from study to study, the group almost invariably consists of young males. The Department has found that the group consisting of males between the ages of 18 and 29 is a particularly good candidate for the appellation "population at risk". The percentage of admissions to DOR represented by this group has been consistently over 50 percent since 1960. Furthermore, that percentage has been growing. While the population at risk accounted for 55.8 percent of admissions in FY 1960-61, it represented 73.5 percent of admissions in FY 1975-76.

Inmate admissions were therefore projected using a multiple regression analysis of the population at risk and the Florida unemployment rate. The population at risk and the unemployment rate for a given calendar year were used to project inmate admissions for the fiscal year beginning that year.

The correlation coefficient for admissions with the population at risk was .92 and for admissions with the unemployment rate was .67. The multiple correlation coefficient for admissions with the population at risk and the unemployment rate was .99.

Three year projections of admissions were based on both the population at risk and the unemployment rate. The regression equation was:

ADM = 12.194*POPRISK+337.4*UNEMP-3928.3

The long range projections of admissions--covering a period of twenty-three-years--were based solely on the population at risk. The regression equation was:

ADM = 14.436*POPRISK-3337.6

The two sets of projected admissions were fed into the release module producing short-term and long-term projections of releases and population.

VIII

SUMMARY

In summary, the Department has expended considerable effort in trying to develop more effective methodologies for the prediction of the future inmate population.

The earliest attempts to make population projections centered around projecting the past percentage increase, and the linear regression method. In the application of linear regression utilized by the Department, the data base was the past inmate population. This particular method has some important advantages. The necessary data is available and is relatively easy to collect. Another advantage of this technique is that reasonably accurate predictions are possible when there has been a stable and linear historical trend, and when that trend is expected to continue. On the other hand, this form of linear regression has serious limitations. The limitations involve the inability of the technique to reveal upcoming major changes in the inmate population level. An example of the strengths and weaknesses of the method is furnished by the experiences of the past several years. Because the pattern of growth had been fairly stable until 1974, projections using linear regression were quite accurate. The projections were very wide of the mark, however, for the period after 1974, when the inmate population suddenly began growing at an unusually rapid pace.

In order to try and overcome the major weaknesses of linear regression, the Department considered using the SIMMODG and SPACE models. The objective of these models was to simulate some of the workings of the criminal justice system. By examining the number of arrests at a given time, as well as the status of the population on probation and parole, it appeared possible to anticipate important changes in the correctional population. This type of prediction model avoids the weakness of linear regression in that it does not automatically assume a continuation of a given historical pattern. On the contrary, it is assumed that a major change at the arrest stage of the criminal justice system will be reflected several months later in the prison population.

The major weakness of these models is the short time frame for which projections are possible. Projections covering an extended period, say three or ten years, are not possible. That is not an acceptable state of affairs for the Department, which must plan for the construction of many new facilities and the resulting capital outlay.

The Department has therefore developed new techniques in order to more accurately predict the future inmate population. The Simulated Losses/Admissions Model breaks down the prediction into releases and admissions. Different methods are used for each. For releases, a simulation model has been used in order to predict a probable release date for those currently in custody, as well as for those not yet admitted to prison. This method gives a far more complete picture of releases than simply recording the gross number of annual releases. In particular, this method makes it possible to study the residual population-that group of inmates serving long sentences that has been building up in the prison system.

For predicting admissions, the Department has employed a multiple regression equation. It has been found that there are strong correlations between the number of inmate admissions and both the population at risk and the state unemployment rate. This method of predicting admissions has clear advantages over the methods previously used by the Department. Unlike the linear regression method, no assumption is made that past trends in inmate population growth will continue indefinitely into the future. On the contrary, inmate admissions are predicted to rise or fall depending on expected changes in the predicting factors. The multiple regression method has a major advantage over the SIMMODG and SPACE models in that predictions can be made for a far longer time span.

At present, the Department would like to simulate the workings of the entire criminal justice system in order to improve still further its ability to predict admissions. The closest approximation to such a simulation appears to have taken place in Maryland. In that state correctional planners analyzed arrests, the probability of varying dispositions following arrests, and the expected changes in the composition of the state's population. The major hindrance to a methodology of this sort in Florida has been the paucity and frequent unreliability of court data. When and if reliable court data becomes available, the Department will utilize it in order to improve the accuracy of predictions. At present, the Department is confident that the current methods used for the prediciton of the future inmate population represent a significant advancement over those employed in the past.

END