Appendix A: Predicting Prison Populations Literature Review

Introduction

Prison population forecasts are essential for prison administrators and policy makers to make management and budget decisions. Prison population forecasts are also significant for legislators to make informed decisions when passing laws that potentially affect prison populations.

The growth of prison populations in the past 30 years has made prison population forecasts necessary. Between 1980 and 1990 the U.S. prison population grew by approximately 134% (U.S. Department of Justice 1995). The prison population increase slowed between 1990 and 2000, but still grew by 69% (U.S. Department of Justice 2001). Martinez (2009) made the argument that prison population forecasts are crucial due to the length of time it takes to build a new prison. After legislators have approved funding for construction of a new prison, it can take two years for a prison to be built and staffed.

Legislative and policy decisions have a direct impact on prison populations. According to a report produced by the Federal Bureau of Investigation in 2004, U.S. crime rates decreased in the previous10 years, but the prison population for that time period increased. The cause of the prison population increase has been attributed in part to changes in sentencing laws, including: longer prison sentences for some crimes; three strikes legislation; stricter habitual offender laws; an increase in mandatory minimum stays; tougher policies imposed on criminals in prison, on parole or probation; and the war on drugs (Martinez, 2009).

Prison Population Forecast Models: Then and Now

Since the 1960s, trying to project future prison populations has proven difficult. In 1984, the Federal Bureau of Prisons (BOP) announced:

> ". . . The 'state of the art' for predicting prison populations is still in its infancy and accurate and reliable methodologies simply do not exist. Our review of numerous prison population projection studies conducted by national experts reveals, with the wisdom of hindsight, that their projections have continually been in error."

In 1984, the General Accounting Office (GAO) surveyed the BOP, the District of Columbia, and the 50 states to find what methods were used to forecast prison populations. The GAO found that states used more than one method to forecast. Fifty-two percent analyzed admissions and releases to forecast prison populations. Nineteen states (38%) used trend analysis based on past prison populations, 17 (34%) performed a simulation of policies and practices then assessed how changes would impact the prison population. Thirteen states (26%) performed linear regressions using factors such as unemployment rates, which seemed to correlate to prison populations when the rates are lagged six months to a year. Twelve states (24%) used multiple linear regression, 20% projected future populations based on design or rated capacity of their facilities. Two states based projections on a "consensus statement" or group opinion (GAO, 1984).

In 2008, the American Correctional Associations in its journal, *Corrections Compendium*, published results of a survey of US and Canadian correctional systems. The agencies were asked to project their populations for the years 2008, 2010 and 2012. The survey found 28 U.S. correctional systems perform internal projections. The systems used a variety of methods including stochastic models, a flow model method pioneered in Texas, autoregression integrated moving average (ARIMA), and a micro-simulation model. Agencies also reported analyzing their own historical population data and conducting a general simulation of admissions, lengths of stay, and departures. If not developed and performed within their systems, the departments identified outside sources such as JFA Associates, the Connecticut Office of Policy and Management, a local university, the Criminal Justice Estimating Conference, and specific state agencies and boards. Twenty-seven agencies reported their figures were considered to be accurate or reasonably so, higher by 5 of the agencies and lower by 7 of the agencies (Corrections Compendium, 2008).

The 2008 Corrections Compendium survey revealed the methodologies used to produce prison population projections have not changed significantly since the GAO's 1984 report. Martinez (2008) stated, ". . .The methodologies used to produce prison population projections have not changed significantly in the past 10 to 15 years, despite the fact that advancing computer technologies could make the task much easier."

In the past it was thought that the total number of citizens in the population primarily affected the prison population. Based on this assumption, prison populations were expected to reach their pinnacle in the 1990s and start their decline with baby boomers passing out of the crime age population (18-36) (Barnett, 1987). As we now know, the rate of growth of prison populations has slowed, proving the inadequacy of predicting prison population growth on the total population of citizens in the community.

Prison population forecast models based on historical population data, admissions, lengths of stay, and departures are limited to the scope of population growth trends and legislation that are current at the time the forecast is run (Barnett, 1987). More advanced models such as the flow, stochastic, autoregression integrated moving average (ARIMA), and micro-simulation models are considered to be more accurate than models based on primarily historical

data and can be adjusted to include changes in policies and practices (Martinez, 2008).

Conclusion

Experts agree that predicting prison population is not an exact science. Predicting prison populations is a combination of facts and probabilities (Martinez, 2009). The state of the art prison population forecast model does not currently exist. The rapid advancement of computer technology should be utilized to produce the state of the art prison population forecast model. Experts believe the state of the art prison population forecasting model should be:

- A computer simulated model (BOP 1984, Martinez 2008)
- Intuitive so those who do not regularly deal in statistical mathematical concepts could understand the prediction output and could input their own queries (Martinez 2008)
- Able to answer 'what if' scenarios to help legislatures make informed decisions when passing laws that affect prison populations (Martinez 2008)
- Capable of taking into account the vast number of variables to produce an accurate forecasting model (BOP 1984, Martinez 2008).

References

American Correctional Association. (2008). Prison Populations. *Corrections Compendium*.

Barnett, A. (1987). Prison Populations: A Projection Model. *Operations Research*, 35(1), 18-34.

Martinez, P. E., (2008). Projecting Prison Populations Starting with Projected Admissions. *The Prison Journal*, 88(4), 493-516.

Martinez, P. E., (2009). Projecting Felony Intakes to the Justice System. *The Prison Journal*, 89(4), 383-400.

New Mexico Sentencing Commission. (2008). Possible Reasons for Decline in New Mexico Corrections Department Inmate Population.

Sabol, WJ., West, H.C., Cooper, M., (2010). Prisoners in 2008. found at http://bjs.ojp.usdoj.gov/index.cfm?ty=pbdetail&iid=1763 and Probation and Parole in the United States, 2008 can be found at http://bjs.ojp.usdoj.gov/index.cfm?ty=pbdetail&iid=1764. (NCJ-228417).

Spelman, William. (2009). Crime, cash, and limited options: Explaining the prison boom. *Criminology & Public Policy*. 8: p.32.

U.S. Department of Justice. (1995). Prisoners in 1994 (Bureau of Justice Statistics Bulletin NCJ151654). Washington, DC: Government Printing Office.

U.S. Department of Justice. (2001). Prisoners in 2000 (Bureau of Justice Statistics Bulletin NCJ188207). Washington, DC: Government Printing Office.

U.S. Department of Justice. (2007). Prisoners in 2006 (Bureau of Justice Statistics Bulletin NCJ205335). Washington, DC: Government Printing Office.

APPENDIX B: DESCRIPTION OF DATA FILES

Admissions File	
Variable	***Definition***
State id number	Unique offender/incarceration identifier
Gender	Sex of offender
Race	Race of offender
Date of birth	Date of Birth of offender dd/mm/yyyy
County of residence n/a	This field is optional
Marital Status	This field is optional
Statute	This field should represent the *most serious offense statute* the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized.
Offense Description	This field should describe the most serious offense the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized and standardized offense name used.
Jail credits	This field should represent the total number if pre-trial/ jail credits to be awarded to the offender.
Admission type	i.e., parole violator technical, parole violator new charge, probation violator technical, probation violator new charge, new court commitment, escapee returned, etc.
Sentence length (Maxdays)	This field should represent the total net sentence the offender will serve under DOC custody. All consecutive and concurrent calculation should be applied. Lifers will also need to be determined from this field.
Parole eligibility date	This field should represent the first date in which an offender is parole eligible. dd/mm/yyyy
Goodtime earning class	This field should represent the number of goodtime days per month the offender is eligible to receive.
Offense Class Code	This field should represent the most serious offense the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized; standardized codes should be employed.
Mandatory release date (flatdate)	This field should represent the absolute latest day the offender will be released. dd/mm/yyyy
Initial classification level	This field should represent the results of the initial classification, i.e. minimum, medium, maximum, close
Final custody level level	This field should represent offender custody level placement after overrides
Projected release date	This field should provide the projected release date assuming all future good-time will be awarded
Offense severity	Severity of current offense
Arrest date	Date of offenders arrest for current offense
Offense date	Date crime offender is currently held for was committed
Sentence date	Date offender was sentenced for most current/serious offense
Sentence Begin date	Sentence begin date
Institution start date	Institution admission date

Confined File	
Variable	***Definition***
State id number	Unique offender/incarceration identifier
Gender	Sex of offender
Race	Race of offender
Date of birth	Date of Birth of offender dd/mm/yyyy
County of residence n/a	This field is optional
Marital Status	This field is optional
Statute	This field should represent the most serious offense statute the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized.
Offense Description	This field should describe the most serious offense the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized and standardized offense name used.
Jail credits	This field should represent the total number if pre-trial credits to be awarded to the offender.
Admission type	i.e., parole violator technical, parole violator new charge, probation violator technical, probation violator new charge, new court commitment, escapee returned, etc.
Sentence length (Maxdays)	This field should represent the total net sentence the offender will serve under DOC custody. All consecutive and concurrent calculations should be applied. Lifers will also need to be determined from this field.
Parole eligibility date	This field should represent the first date in which an offender is parole eligible. dd/mm/yyyy
Goodtime earning class	This field should represent the number of goodtime days per month the offender is eligible to receive.
Offense Class Code	This field should represent the most serious offense the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized; standardized codes should be employed.
Mandatory release date (flatdate)	This field should represent the absolute latest day the offender will be released. dd/mm/yyyy
Current classification level (1-6)	This field should represent the current classification level of the offender.
Final custody level	This field should represent offender custody level placement after overrides
Projected release date	This field should provide the projected release date assuming all future good-time will be awarded
Offense severity	Severity of current offense
Arrest date	Date of offenders arrest for current offense
Offense date	Date crime offender is currently held for was committed
Sentence date	Date offender was sentenced for most current/serious offense
Begin date	Sentence begin date
Institution start date	Institution admission date

Release File	
Variable	***Definition***
State id number	Unique offender/incarceration identifier
Gender	Sex of offender
Race	Race of offender
Date of birth	Date of Birth of offender dd/mm/yyyy
County of residence n/a	This field is optional
Marital Status	This field is optional
Statute	This field should represent the *most serious offense statute* the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized.
Offense Description	This field should describe the most serious offense the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized and standardized offense name used.
Jail credits	This field should represent the total number if pre-trial credits to be awarded to the offender
Admission type	i.e., parole violator technical, parole violator new charge, probation violator technical, probation violator new charge, new court commitment, escapee returned, etc.
Sentence length	This field should represent the total net sentence the offender will serve under DOC custody. All consecutive and concurrent calculations should be applied. Lifers will also need to be determined from this field.
Parole eligibility date	This field should represent the first date in which an offender is parole eligible. dd/mm/yyyy
Offense Class Code	This field should represent the most serious offense the offender is currently serving, even if it is not his/her longest sentence. DOC established hierarchy of offenses should be utilized; standardized codes should be employed.
Mandatory release date	This field should represent the absolute latest day the offender will be released. dd/mm/yyyy.. but this is as of the date of release
Release date	This field should represent the actual date the offender was released from DOC custody.
Release type	This field should represent the reason for an offender's release, i.e., parole, discharged, escape, transfer to another state, etc.
Total statutory monthly merit time earned	This field should represent the total merit time credits an offender received during his/her stay at DOC.
Total goodtime credits lost	This field should represent the total credits an offender lost due to disciplinary infractions during his/her stay at DOC.
Total goodtime credit forfeited	This field should represent the total goodtime credit forfeited by an offender during his/her stay at DOC.
Total goodtime credit restored	This field should represent the total goodtime credit restored to an offender during his/her stay at DOC.
Total other (lumpsum) credits	This field should represent the total 'other' credits an offender received during his/her stay at DOC (including credits for education, work, etc.).
Finial classification level (1-6)	This field should represent the last classification level the offender was in before release, i.e. minimum, medium, maximum, close
Final custody level	This field should represent offender custody level placement after overrides
Projected release date	This field should provide the projected release date assuming all future good-time will be awarded
Offense severity	Severity of current offense
Arrest date	Date of offenders arrest for current offense
Offense date	Date crime offender is currently held for was committed
Sentence date	Date offender was sentenced for most current/serious offense
Begin date	Sentence begin date
Institution start date	Institution admission date

Goodtime Release File	
Variable	***Definition***
State id number	Unique offender/incarceration identifier
Lump Sum Total	Total amount of times in days an offender was awarded
Lump Sum Comments	Comments relating to the lump sum award: comments are in a free text field and will indicate reason for award.

APPENDIX C: METHODOLOGY

The prison population time series forecasts used to produce this report are based on observed prison population data. It is understood that there are many factors that drive prison populations, including demographic trends, arrest rates, the number of criminal cases filed in district court, conviction rates, the availability of diversion programs, sentence lengths, admission rates and release rates, availability of earned meritorious deductions and parole readiness. The observed prison population is a result of all those factors and others. When new laws or polices come to bear which significantly affect the prison population, it is recommended that a new long-term forecast be produced which incorporates new data that reflects the changes.

Time series forecasting consists of examining historical prison population data, identifying potential methods for the forecast, fitting the data to a model which will use the data to produce a forecast into the future, and then testing the model. Testing includes assessing the overall model fit, producing estimates and comparing those estimates to actual data to see how well the chosen model performs. Diagnostic checks are applied to the differences between the estimated and actual counts to ensure that the model adequately explains and extracts all information that the historical data has to offer. It may turn out that more than one model specification fits the data well. When choosing between different candidate models, there are fit statistics produced for each model that can be compared.

The methodology described above was augmented at various steps by conversations with colleagues who have historical knowledge regarding prison population trends, factors that drive population and insight into population patterns. Moreover, Sentencing Commission staff held quarterly meetings with New Mexico Corrections Department staff to discuss inmate population trends. This information was crucial for choosing the starting date from which to forecast for males and females, respectively.

Next, examination of the daily and monthly high counts for males and then females was conducted via graphical analysis of the historical data plotted against time. As a result of this analysis, we came to the two following conclusions: 1) that the men's and women's population should be modeled separately and 2) that using monthly high population counts would be the best way to proceed.

Working with the male and female population time series data separately, we moved from graphical analysis to fitting and diagnosing models. It became apparent that each time series called for a different methodology in order to produce the forecasts. For the males, an Exponential Smoothing (ES) model was used and for the females the Box Jenkins (BJ) method was used to specify an Autoregressive Integrated Moving Average (ARIMA) model. Each of these methods are discussed below in the male and female sections.

MALES

The historical monthly high data for males included the time range between April, 2004 through March, 2012. The starting date was chosen after initial examination of the historical data, discussions among staff and then performing model fitting and diagnostics. It was found that the Exponential Smoothing method was best suited to handle the male data. Specifically, we tested a Winter's Additive (WA) model using a one period backward lagged dependent variable. The WA has an ARIMA equivalent or is a special case of such. For the ES method, the forecasts are based on weighted averages where the future values are weighted averages of past population observations, with more recent observations given more weight in the forecast than population observations in the more distant past.

The WA model performed better than other ES model candidates. As opposed to the ARIMA model, the residual diagnostics were very good implying that this model specification adequately explained the data process for the time period used. This model captured a slowly changing seasonal pattern that exhibits constant or additive seasonal variation along with a slowly changing linear trend. As apparent in the forecast, the varying cycle repeats in an upward trend.

Since ES methods are not based on a formal statistical method, it is recommended that a back forecast be produced and checked for accuracy. In this case, the data range was cut off at February 2011 and a forecast for the period between March 2011 and March 2012 was produced. The forecasted monthly highs were compared against the actual male population via calculation of the percentage difference between the two. The forecasted values were slightly lower, with an average difference over the 13 months of 1.16%. The highest differences were present in August, September and October of 2011 and the lowest differences were present in March and April of 2011 and March of 2012.

FEMALES

The historical monthly high data for females includes the time range between July 2010 through March 2012. The starting date was chosen after performing graphical analysis and conversations with colleagues regarding

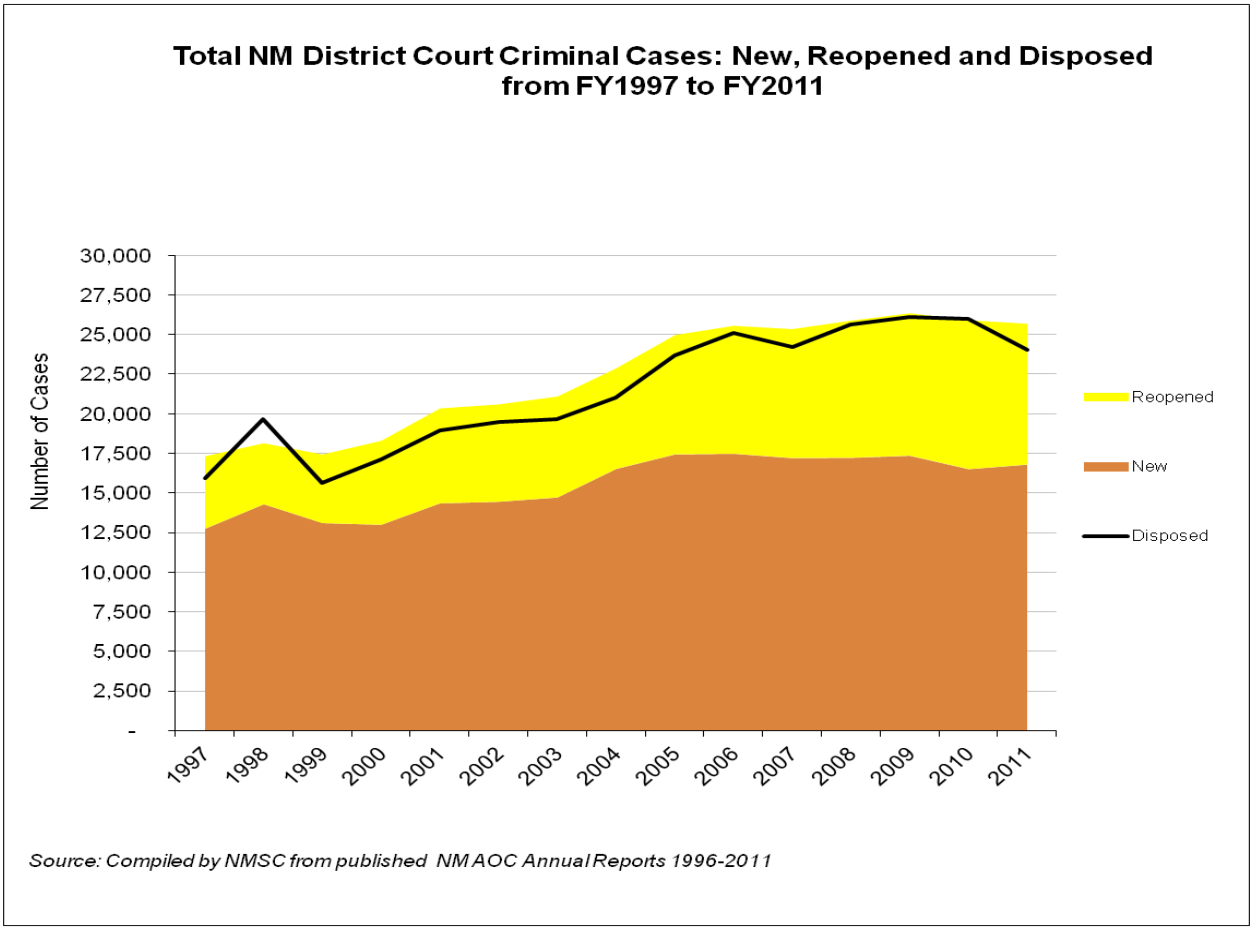
recent history specific to the female population. The information regarding recent history was important in choosing a time frame in which the population could be expected to exhibit a relatively stable pattern.

Choosing an appropriate forecasting model for the women entailed utilizing the Box Jenkins method to specify an ARIMA model. The Exponential Smoothing method did not adequately describe the female population data. The primary difference in the methodology is that the auto and partial autocorrelation functions (ACF's and PACF's) are also examined graphically to identify potential models. These show how correlated each value is with its past value for a number of periods in the past. They also aid in ARIMA model identification, including whether a difference is needed to account for non-random patterns in the data, such as seasonal effects.

Specification of the forecasting model for the female population was a two-step process. First, the data was fit to a seasonal ARIMA model. It was found to follow an autoregressive (AR) and a seasonal AR of order one. This model (Model I) performed well for a short term forecast. However, examination of the ten year forecast revealed problems, attributable to the fact that with so few observations it is difficult to capture long-term patterns.

Consequently, Model I was then used to forecast out one year, thereby providing 12 more observations. The next step was then to repeat the model fitting process for the time range of July 2010 through March 2013. For the last twelve months of this range, the observations are actually forecasted values from Model I. The results from re-fitting the data produced a similar model with the exception that the AR process was of order two, and a first difference was used. The Box Jenkins method was implemented when specifying both Model I and II, inclusive of fit assessment and residual checks. Model I fit the data well for the shorter time period, while Model II performed well for the second time period. As with the men, the women's long term forecast exhibits varying seasonality following an upward trend.

APPENDIX D: NEW MEXICO JUDICIARY DATA



Source: Compiled by NMSC from published NM AOC Annual Reports 1996-2011

New Mexico District Court Criminal Cases FY1997 to FY2011				
Year	New Cases	Reopened	New + Reopened	Total Disposed
1997	12,743	4,570	17,313	15,905
1998	14,290	3,848	18,138	19,635
1999	13,101	4,327	17,428	15,625
2000	12,995	5,300	18,295	17,119
2001	14,349	5,991	20,340	18,972
2002	14,449	6,141	20,590	19,453
2003	14,718	6,372	21,090	19,660
2004	16,522	6,349	22,871	21,007
2005	17,439	7,530	24,969	23,708
2006	17,482	8,071	25,553	25,083
2007	17,206	8,139	25,345	24,224
2008	17,226	8,657	25,883	25,648
2009	17,359	8,983	26,342	26,111
2010	16,509	9,396	25,905	25,963
2011	16,796	8,888	25,684	24,018