Document Title:	Crime-Control Effect of Incarceration: Reconsidering the Evidence, Final Report
Author(s):	Bert Useem ; Anne Morrison Piehl ; Raymond V. Liedka
Document No.:	188265
Date Received:	06/18/2001
Award Number:	98-IJ-CX-0085

188265

The Crime-Control Effect of Incarceration: Reconsidering the Evidence

Final Report to the National Institute of Justice (98-IJ-CX-0085)

Bert Useem
Department of Sociology, University of New Mexico

Anne Morrison Piehl
John F. Kennedy School of Government, Harvard University

Raymond V. Liedka
Department of Sociology, University of New Mexico

November 1999, revised January 2001

Points of view in this document are those of the authors and do not necessarily represent the position of the U.S. Department of Justice.

Project Summary

The Crime-Control Effect of Incarceration: Reconsidering the Evidence

Over the last quarter century, the population behind bars has grown at an exponential rate. From 200,000 inmates in 1973, the state and federal and state inmate population surpassed the one million mark in 1994, with more recent growth bringing the total to 1.2 million. This build up has been costly, in the first place to taxpayers but also for those behind bars, who might otherwise be in community sanctions programs. Perhaps the money and human toll are worth it, if prison expansion drives down the crime rate. Yet we do not know this. During the middle stages of the build up, many criminologists dismissed the prison expansion as a failure because rates of imprisonment were rising, but crime rates were steady. If prison works, why no reduction in crime? More recently, the crime rate has fallen, and it has become common to explain the decline in part by prison expansion. Perhaps, but the point remains largely speculative. Other factors may have caused the decline, with high rates of imprisonment contributing little or nothing.

Researchers disagree about whether changing the rate of imprisonment has any effect on the crime rate. The arguments that a larger prison population generates crime are perhaps less common than those for the opposite claim, but are wholly plausible. Using very different empirical methods, recent studies by Levitt (1996) and Marvell and Moody (1994) arrived at similar conclusions about the effect of prison space on crime rates: about 15 Uniform Crime Reports index crimes averted for every additional inmate. The results of these studies have been used to support expanding the use of incarceration. Given the policy importance of such findings, these studies deserve close scrutiny. Furthermore, it is essential to continually update our understanding of empirical relationships as further experience unfolds. (Our companion paper includes a review of a number of other studies.)

This paper argues that, due to the sources of variation used to estimate this effect, the methodologies used in each of these studies are such that the findings are of limited value for thinking about punishment policy. We offer detailed critiques of these two studies, more general critiques of the broader literature concerned with the measurement of the empirical relationship between incarceration and crime, and our own estimates of the relationship. In our estimates, we extend analyses to use a flow measure of prison admissions constructed from a data source which has gone unexploited in the literature. The results are extremely sensitive to modest changes in specification, which strengthens our conclusion that much more theoretical and empirical work must be done before researchers can be confident we know the true magnitude of the relationship of interest.

The most difficult problem in estimating the effect of prison on crime is that causation may flow in both directions. Increasing imprisonment may reduce the level of crime, but

increasing crime will put more people in prison, other factors held constant. One approach to break this "simultaneity" problem is to use instrumental variables to estimate the effects of imprisonment on crime. In this context, instrumental variables are constructs that affect prison populations but are not directly related to crime. Steven Levitt (1996) claims to have discovered one such instrument: prison overcrowding litigation involving an entire prison system. This insight forms the basis for his influential paper.

In attempting to solve the simultaneity problem by using instrumental variables regression, the choice of a proper instrument is crucial. There are three problems with Levitt's choice. First, Levitt's instrument may be as susceptible to the simultaneity problem as the original measure of prison population. It is difficult to see how, if rising crime rates cause increases in prison populations, they do not also play a role in determining overcrowding litigation. In these cases, the instrument is correlated with the outcome of interest and therefore does not solve the simultaneity problem. Second, for the period under study, only twelve states experienced system-wide litigation. A lot is riding, then, on whether the experience of the twelve states at a particular time can be generalized to the other 38 states, or even the same states at a different time. One would gain confidence if the 12 states are roughly similar to the remaining 38 states. But that was not the case. The majority of the states, seven of the 12, are Southern, with the remaining five each small states. Levitt (1996: 326) is correct to point out that Southern states historically had high rates of incarceration, and this helps explain why Southern states were over-represented among those subjected to overcrowding litigation. The problem is that Southern correctional exceptionalism runs much deeper than this. One has to wonder whether the results would have been different had there been instrumental variables available for the large northern and western states. We are left unsure whether the same effects occur in other contexts, which is hardly a secure basis from which to set policy with such serious consequences for the American public.

Perhaps no study has had a larger impact on the prison/crime issue than a paper published in the Journal of Quantitative Criminology by Thomas Marvell and Carlisle Moody (1994). The authors regressed crime rates on prison population over a 19-year period, finding that the size of the state prison population had a significant, short-term negative impact on crime. Expressed in the metric of crimes averted, "each additional state prisoner averted at least 17 index crimes on average, mostly larcenies" (1994:136). Like Levitt, Marvell and Moody regressed crime on prison. Their approach to the endogeneity issue is to examine the data for evidence of reverse causality, that is, crime driving incarceration. To do this, they employ the technique developed by Granger (1969) for analyzing causal ordering using time series data. Marvell and Moody conclude that a model regressing crime on prison populations is well specified for determining causal relationships, at least for short-run movements in the series.

There are several technical problems with Marvell and Moody's approach to the prison/crime relationship. First, they use a very narrow statistical test to justify their empirical model, and they implement that test incorrectly. (In particular, they fail to consider auxiliary variables in the Granger causality test.) The results of this test lead them to shift their attention from the long run to the short-run effects of SPP on crime. We believe that the decision by

Marvell and Moody to ignore the long-term relationship between state prison population (SPP) and crime is both unwise and unnecessary. It is unwise because most observers are interested in the long-run relationship between prison population and crime rates. While information on short run fluctuations has its own intrinsic value, it is the long-run impact of SPP on crime that grips us, both theoretically and for policy considerations.

It is unnecessary because there are other techniques for capturing complicated relationships among a set of variables. Additionally, the Marvell and Moody model is not fully specified. In particular, they use a very narrow set of control variables. A more complete specification would include (in addition to the population variables they considered) two additional types of control variables: (a) states' economic and social characteristics, and (b) alternative sanctions, that is, information on the number of probationers and parolees. With regard to the error structure, Marvel and Moody include a lagged dependent variable on the right-hand side of their model, ostensibly to control for dynamic effects and autocorrelation. In our view, the problem of autocorrelation can be more effectively handled by directly including an autoregressive term into the error structure of the model.

When we investigate the impact of making modest improvements to the M&M approach, we find that the results are extremely sensitive to specification. There are two types of modifications we make: extending the set of control variables to account for the influences that economic and socio-demographic factors may have on crime rates and extending the time period analyzed. Our controls include, in addition to the measures of the age distribution of the population used by Marvell and Moody, two economic variables (the state unemployment rate and a measure of wages for the lower end of the wage distribution), the percentage of the state population African-American (African-Americans are over-represented among both perpetrators and victims of crime), the percentage of the population residing in metropolitan areas (urban areas have higher levels of crime, possibly due to the concentration of viable crime targets in a small area). In some portions of the analysis we include three additional variables capturing the use of sanctions that are alternatives to incarceration: probation, parole, and jail.

When we variously add these control variables or add the years 1990-1997 to the specifications used by Marvell and Moody, the estimated effect of prison population on crime rates is disturbingly unstable. Depending on the time frame covered in an analysis, or on the set of control variables included in an analysis, the estimated effect of prison population was significantly positive *or* significantly negative. This volatility in the estimated effect of prison on crime rates is due only to the length of the time period covered and the presence or absence of an extended set of control variables neglected by Marvell and Moody. The robustness of an estimated negative coefficient for prison in the Marvell and Moody analysis is now under some question. As a result, the additional confidence from multiple studies with similar results is also under question.

We also conducted a number of analyses using specifications which seemed much more sensible to us. Each of the following changes in specification change the nature of the estimated effect of incarceration: the inclusion of fixed effects (as opposed to random effects), the use of an autoregressive error structure (or not), or using the natural logarithm of prison population (rather than levels). In some specifications, the effect of prison was significant and negative. In one the effect was significant and positive. In still others, prison population had no statistically significant effect at all. This behavior is in strong contrast to the robustness of the estimates of the control variable. The magnitude and significance of the effects of control variables rarely changed from model to model. That the effect of prison population *did* change from model to model compels researchers to rethink the specification of their models.

Given that it seemed to us that the question of the effect of incarceration on crime was not settled, we decided to consider a broader framework. In particular, while continuing to work with a state-level panel analysis, we exploited data from the National Corrections Reporting Program (NCRP) to posit an additional set of relationships. The NCRP collects data on every admission and release from prison for those states that report. The program began in 1983, and the most recently available data are from 1996. For each admission, demographic information (age, sex, race, and ethnicity), prior incarceration history (prior jail time, prior prison time), and current offense and sentence information are recorded.

There are three key advantages to the NCRP data. First, they allow us to analyze the flow of inmates into prison instead of (or in addition to) the stock measure of prison population generally used in this literature. This may be a very enlightening improvement to the model, as the data provide a flow measure (admissions) to explain a flow measure (crime rates). Whether flow matters more than stock in crime reduction is an empirical question. For example, assume for a moment that prison reduces crime primarily through deterrence. The question would remain open whether potential offenders are more likely to be dissuaded from crime by the total number prisoners in a state (as measured by stock) or the total number of offenders the state recently sentenced to prison (as measured by flow).

Second, the NCRP data provide information on each prisoner's incarcerating offense. Of particular interest is to separate drug offenders from other felony admissions, and then to determine whether imprisoning drug offenders has the same effect on crime as imprisoning non-drug offenders. Third, these data allow us to construct a measure of sentencing practices which, in turn, permits us to determine the effect of incarceration net of the severity of sentencing. At least in theory, we can then separate out the impact of an increase in prison population arising from the widening of the net from an increase due to longer sentences to the same types of offenders as incarcerated previously.

Aggregate state-level counts of the number of admissions are created from the master data files by summing the number of cases (admissions) by reporting state and reporting year. Likewise, data on the most serious offense committed – identified as the offense with the longest sentence – are used to classify every prison admission by offense type (in particular, drug vs. property or violent). It is important to realize that the categorization of offense types is not obvious or simple. Our definition – a drug offender is one whose most serious offense was a drug offense – may seem straightforward and reasonable. However, alternative definitions of "drug offender" could include everyone admitted to prison with any drug offense or could distinguish between possession and trafficking.

Given our definition, the number of admissions by offense type for every state-year is then determined in the same manner as total admissions. There is some loss in state-year level data at this point, since offense data are missing for some states in some years. For instance, Arkansas in 1990 and 1991 has no offense information for 32% and 43% of admissions, respectively. For any analysis involving total admissions, we include these two years. When the analysis breaks out admissions by offense type, we drop observations with a high proportion of missing data, because the missing offense data would lead to a grossly underestimated aggregate count of admissions and other sorts of potential bias.

Sentencing practices are of interest in our study because prison populations can expand in two ways. A larger proportion of convicted offenders may be sent to prison or the sentences given out may be increased in length. In fact, it may be that sentencing practices are the driving force explaining variation in crime rates. Under this perspective, incarceration may be an imperfect and incomplete proxy for sentencing; not accounting for sentencing could lead to spurious effects of incarceration. Sentencing -- as a manipulable policy alternative akin to probation -- is of interest on its own merits and not merely as a control variable. However, in the current research we will be primarily interested in sentencing as a control.

The results of these new analyses do not clear the water which was muddied by the careful examination of the Marvell and Moody analyses. However, there are a few provocative findings. As before, the effect of prison population is *not* robust with respect to the specification of the model. Although the stock and flow measures of prison inmates are highly correlated, they clearly are not that similar when used to model crime rates. For the stock measure, state prison population was significant in several of the models (albeit with a direction that bounced around) whereas the flow measure of prison admissions never achieved statistical significance. While we found no clear definitive model for either, we do observe that they perform differently even in models with the same specification. Hence, it is important to continue research using flow, as its role is no less uncertain than the traditional measure of the prison population.

A separation of admissions into non-drug and drug offender admissions points to a mixed bag of effects, similar to that observed in the prison stock analysis of the earlier section: drug offender admissions sometimes have significant positive effects on crime rates and sometimes have significant negative effects on crime rates. Model specification would seem to be an important component to the kind of results observed. Our measure of sentencing practices does not seem to matter, although the roughness of the measure may be the reason.

In a final effort to discern the impact of non-drug and drug offender admissions, we analyze the effects of such admissions separately on violent and property crime rates. One possibility is that the effects of non-drug and drug offenders have been masked by consideration of overall index crime rates. More specifically, we might speculate that drug offender admissions would impact property crime rates more than on violent crime rates. Separation of the crime rate into violent and property crime rates does not clarify the issue so much. The only discernable effect from this set of analyses is a positive effect of drug offender admissions on property crime rates in a fixed-effects model.

Many now argue that the United States relies excessively on prisons. Others are heartened by the recent decline in the crime rate, and want to credit prison expansion for at least some of this trend. In our view, both positions are reasonable, and unproven. One purpose of this report has been to shake the reader's confidence in existing estimates. We have not provided a "new and improved" estimate, though we think that this is entirely possible and is the goal of our future work.

We lack confidence in existing efforts for several reasons. The key one is the lack of robustness of the regression results. If indeed there were a deep, causal association between prison population and crime rates, minor changes in specification or the years covered would alter only slightly the results. In conducting our own analyses, we did not find this. The significance levels changed, and even the signs reversed, depending time period covered, control variables included, and the estimating techniques used. We had no way to explain these perturbations, and the only reasonable statement at this stage is that we are not sure what is going on. It would be hard for a researcher to argue that he or she had identified *the* right model from among the candidates.

Finally, it quite possible that other issues raised here, such as the use of other correctional sanctions and the possibility of nonlinear impacts, matter a lot. Additionally, a different picture sometimes seems to emerge if one uses prison admission flows rather than prison inmate stock. One issue, with clear policy implications, is whether the imprisonment of a large number of drug offenders has any impact on the crime rate. There is some evidence that the effect of non-drug admissions is not the same as that of drug offender admissions. The difficulty in making a definitive conclusion comes from the inability of statistical devices to make choices amongst different model specifications, and the instability in estimated effects across those different model specifications. Future work must begin to pay real attention to the choice of model specification, and find justification for chosen specifications in theoretical considerations.

The Crime-Control Effect of Incarceration: Reconsidering the Evidence

Over the last quarter century, the population behind bars has grown at an exponential rate. From 200,000 inmates in 1973, the state and federal and state inmate population surpassed the one million mark in 1994, with more recent growth bringing the total to 1.2 million. This build up has been costly, in the first place to taxpayers. Between 1980 and 1993, government spending on corrections increased by 363 percent, from $6.9 billion to $31.9 billion (Maguire and Pastore 1997, p. 3). For those behind bars, who might otherwise be in community sanctions programs, the costs must be calculated, not only in the pains of imprisonment endured, but reduced lifetime earnings and broken families. For society as a whole, high rates of imprisonment tread upon the value of keeping parsimonious the state's role in the lives of the citizenry. Prison condemns behavior through the greatest intrusion possible of government into citizens' lives – the deprivation of individual liberty. If nothing else, more prisons mean more government.[1]

Perhaps the money and human toll are worth it, if prison expansion drives down the crime rate. Yet we do not know this. During the middle stages of the build up, many criminologists dismissed the prison expansion as a failure because rates of imprisonment were rising, but crime rates were steady. If prison works, why no reduction in crime?[2] More recently, the crime rate has fallen, and it has become common to explain the decline in part by prison expansion. Perhaps, but the point remains largely speculative. Other factors may have caused the decline, with high rates of imprisonment contributing little or nothing. For example, there may have been a cultural shift in this period, perhaps toward greater respect for authority and disrespect for those who violate the law. This cultural shift may have produced both lower crime rates and higher imprisonment rates, with no direct causal connection between prison and crime. In other words, if more and more people see criminals as bad people, fewer would be willing to become criminals and minorities favoring more punishment may become majorities.

The purpose of this research is to come to a better understanding of the effect of imprisoning criminal offenders on the crime rate. The report is divided into four sections. In the first, we discuss the existing literature on the empirical link between prison rates and crime rates. We focus on several key studies, and begin to lay out one of our central points: the genuine limits of what we know. At the same time, the work reviewed provides a point of departure for our own efforts. The second section uses state level panel data to come up with new estimates of the effects of prison on crime. The third section does the same, but here using data collected by the National Corrections Reporting Program on individual offenders entering into and exiting

---

[1]See Moore and Piehl (1998).

[2]This question was asked by Jerome Skolnick (1995, p. 8) in his 1994 presidential address to the American Society of Criminology. Even quite recently, Marc Mauer (1999, p 192) argues that the prison build up has had "relatively little" impact on the crime rate.

prisons in approximately 38 states (the number reporting each year varies). In the last section, we draw conclusions and policy implications. Here we argue the need to leave no stone unturned in studying the effect of prison on crime, and that there are more stones out there.

I. APPROACHES TO THE STUDY OF EFFECTS OF PRISON ON CRIME

Researchers disagree about whether changing the rate of imprisonment has any effect on the crime rate. The arguments that a larger prison population generates crime are perhaps less common than those for the opposite claim, but are wholly plausible. Two broad sorts of points are made in the literature. One concerns the negative impact of imprisonment on individuals and on communities (e.g., Clear 1996; Kahan 1997; Meares 1998; Sampson and Laub 1995). High rates of imprisonment are said to break down the social networks that guide individuals away from crime; remove adults who would otherwise nurture children and mentor youth; deprive communities of income (both licitly and illicitly derived); stigmatize whole groups of people; disenfranchise a significant proportion of inner-city communities; and engender a deep resentment toward the legal system. Also, there has been a fair amount of discussion of the negative impact of imprisonment on offenders' future income and employment-stability (Freeman 1995; Nagin and Waldfogel 1995; Waldfogel 1994; Western and Beckett 1999). Finally, the point is often made that the imprisonment of low-level drug offenders in particular, not only damages their future job-prospects, but immerses such individuals in milieus that are "schools for crime" (Donziger 1996). These individuals, upon release ("graduation"), spread the lessons they have learned to the communities to which they return.

A second sort of crime-amplification argument rests on the idea that expanding the prison system does little more than increase the number of individuals in the criminal-justice net. Zimring and Hawkins (1997:17-19) argue that U.S. prisons did not need to expand to house the most serious, violent offenders — they were already incarcerated. The prison expansion of the last two decades resulted in the imprisonment of large numbers of nonviolent, "marginal" offenders. Ekland-Olson and colleagues (Ekland-Olson et al 1993; Joo et al 1995) contend that a "replacement effect" may be operating, in which vacancies resulting from incarceration are quickly filled by new offenders. Thus, high rates of incarceration may yield a larger offender population — both those in prison and those "recruited" to take their place. The end product is "a larger, more experienced criminal 'work force' and ironically a heightened collective potential for crime" (Joo et al., 1995: 407). Katyal (1997: 2429) takes this argument one step further, advancing the idea that the replacement effects may be at a level higher than 1:1. Incarcerating some drug offenders may merely increase the price of drugs and hence returns, with a net effect of a "further increase in crime."[3]

---

[3]If the idea of "widening the net" is put together with incarceration harming the legal sector prospects for released felons, one could argue that a short-run 1-for-1 replacement effect could be larger than 1-for-1 over time, as the "original" offenders are released and re-enter the illegal sector.

The arguments that incarceration has a crime-reducing effect are somewhat more familiar, if only because they have been employed to defend the recent build up in incarceration. Nevertheless, they have a solid intellectual foundation. Since Becker (1968), economists have argued that if you raise the cost of crime by imposing more or longer sentences, you will see less of it. Crime reduction here occurs through the mechanism of deterrence. The other key argument is that crime is reduced through incapacitation (Freeman 1996; Spelman 1994). This is based on the notion that an offender behind bars cannot commit new crimes against society, unless and until he or she returns to the streets. Therefore, unless there is full replacement, incapacitation will reduce crime over what it would otherwise have been.

In sum, for every argument as to why incarceration should reduce crime, there is another one for why it should generate it. A policy of high rates of imprisonment may "break the back" of crime, or it may backfire. The issue cannot be settled through debate, logic, or appeal to reason. Rather, a research design is needed such that the evidence collected will actually decide the issue.

Broadly speaking, work on the effects of prison on the crime rates divides into two branches. One branch relies on data collected from surveys of inmates, in which inmates are asked about the crimes they committed before their incarceration. Researchers then address policy issues concerning the number of "crimes averted" through possible changes in the level of imprisonment. The original survey work was done at the Rand Corporation (Chaiken & Chaiken 1982). Important analyses using these data include those by Peter Greenwood (1987) and Edwin Zedlewski (1987); other researchers have developed the approach (Spelman 1994), some using new data (Piehl & DiIulio 1995; Piehl, Useem & DiIulio 1999). Still others have challenged the approach (Zimring & Hawkins 1988). Certainly this approach is useful for profiling who goes to, and stays in, prison at a particular time. One gains a much better idea of what prisoners look like, as well as the kinds of crimes they might commit if on the streets. In other words, the approach seems to capture the effects of prison incapacitation on crime.

The biggest drawback of this approach is that it misses entirely prison's potential deterrent effect. The "economic" approach to crime, in which individuals decide to commit crime based in part on the costs from law enforcement efforts, is all but assumed away. If more prison deters, one could not find this out using an approach based on inmate survey data.

The other branch to studying the effects of prison on crime is some sort of regression approach, in which one looks to see if changes in prison population covary with crime, net other factors. There are two principal advantages to a regression approach over the survey approach. First, the survey approach implicitly asks the question, what would happen if those behind bars were released? A regression approach deals with actual changes in the world. Second, a regression approach captures the full impact of prison on crime, whether deterrence, incapacitation, or the many other potential mechanisms described above. If prison generates crime, the regression approach should tell us that, whereas a survey method would probably not.[4]

There are three analytic methods for estimating the covariance: national time- series analysis, instrumental approach, and state-level panel approach. Interestingly, the three best studies exemplifying each of the three approaches yield remarkably similar estimates. This triangulation has not been lost on the authors of these studies, suggesting to them that the results are robust across studies because they are each seeing the same underlying reality. (If you weigh an object with three *types* of scales, similar results give you confidence that you know the object's true weight.) We focus our discussion on these three studies. First, to anticipate, the three studies come to a common conclusion that the elasticity of the crime with regard to the prison rate is about .30. That is, a 10 percent increase in the prison rate will result in a 3 percent drop in the crime rate.

A. National Time Series

National time series analysis is based on the straightforward idea that if you increase the national prison population at time t, you should expect the national crime rate to be affected at time t+1 if, indeed, there is an effect. If the idea is straightforward, estimating models is not.

The most sophisticated recent effort is by economists Robert Witt and Anne Witte (hereinafter, W&W). They regress the log of the Uniform Crime Reports crime index from 1960 to 1996 on the number of federal and state prisoners serving sentences of at least one year (per 1000 population). They include in their model a third variable, increases in labor force participation of women, because (a) it serves as a proxy for a host of social and economic changes that have occurred in families and communities over the last three decades, and (b) the increase of women into the market purportedly increases the opportunity for crime and deepens the supply for potential criminals. (W&W never go very far in specifying these causal links. Also, they do not include other controls in order to isolate the particular contributions of female labor supply.) W&W conclude that the long-run relationship between imprisonment and crime is on the order of a 10 percent increase in the prison population leading to a 3 percent decrease in crime. Of particular interest is that the short-run relationship they estimate is quite similar in magnitude (the point estimate is somewhat smaller).

The W&W time series analysis falls short of a decisive test. The reason is that there were many changes going on during this historic period, and it may be that some other unmeasured variables may have produced the cumulative changes in the crime rates. If correlated with these

[4]It would be possible to ask inmates questions that might allow one to detect crime-generative processes other than incapacitation. For example, if one developed a measure of commitment to criminal values, and found that inmates became more committed to criminal values the longer they were in prison (net other factors), this would be evidence for a "school for crime" effect. No one has done this.

unmeasured variables, prison population may be serving as proxy for them, gaining a level of explanatory credit when it may deserve less (or possibly more). One could collect information on other factors and include them in the model. As we will see, however, panel data has a key advantage over this approach: with repeated observations over a number of units, one is in a stronger position to control for the effects of missing variables.

B. Instrumental Variable Approach

The most difficult problem in estimating the effect of prison on crime is that causation may flow in both directions. Increasing imprisonment may reduce the level of crime, but increasing crime will put more people in prison, other factors held constant. This is why, for example, one is left unpersuaded by Jerome Skolnick (1995, p. 8), when he quotes approvingly the comment: "If the imprisonment/crime reduction hypothesis were valid, the safest jurisdiction in the country would be Washington, D.C., which has by far the highest imprisonment rate in the world." Skolnick of course fails to consider reverse causation: Washington, D.C. has a high rate of imprisonment because of a high crime rate. It remains entirely plausible that had the city chosen an average imprisonment rate, anarchy would have been unleashed on the streets, the crime rate would have been unaffected, or something in between. That no single outcome seems obvious in this thought experiment suggests that more than a thought experiment is needed.

One approach to break this "simultaneity" problem is to use instrumental variables to estimate the effects of imprisonment on crime. In this context, instrumental variables are constructs that affect prison populations but are not directly related to crime. Imagine two pool balls colliding but more or less moving in tandem. At some point, one of the balls hits a pebble on the table — one can than observe whether that externally caused shift in one ball produces a shift in the trajectory of the other ball. In the present context, the idea is to find a variable that causes prison population to change but is otherwise exogenous (unmoved by) the other variables in the equation. We can then see if this external shock to prison population has a corresponding effect on the crime rate. The only difficult aspect of this approach is to locate an appropriate instrument, the right sort of pebble.

Steven Levitt (1996) claims to have discovered one such instrument: prison overcrowding litigation involving an entire prison system. This type of litigation, Levitt shows empirically, has a short-term negative impact on prison population growth. Using this instrument and state-level panel data for 1971-1993, he calculated that each inmate released because of crowding litigation resulted in 15 additional index crimes per year. Levitt conducted his analysis with unusual care and imagination, and the study has gained (much deserved) fame. Still, the "one-more-inmate, fifteen-fewer-crimes" conclusion needs to be treated with caution.

In attempting to solve the simultaneity problem by using instrumental variables regression, the choice of a proper instrument is crucial. There are three problems with Levitt's choice. First, Levitt's instrument may be as susceptible to the simultaneity problem as the original measure of prison population. It is difficult to see how, if rising crime rates cause increases in prison populations, they do not also play a role in determining overcrowding litigation. In these cases, the instrument is correlated with the outcome of interest and therefore does not solve the simultaneity problem.

Second, for the period under study, only twelve states experienced system-wide litigation. A lot is riding, then, on whether the experience of the twelve states at a particular time can be generalized to the other 38 states, or even the same states at a different time. One would gain confidence if the 12 states are roughly similar to the remaining 38 states. But that was not the case. The majority of the states, seven of the 12, are Southern, with the remaining five each small states (e.g., Rhode Island, with only one major prison). Levitt (1996: 326) is correct to point out that Southern states historically had high rates of incarceration, and this helps explain why Southern states were over-represented among those subjected to overcrowding litigation. The problem is that Southern correctional exceptionalism runs much deeper than this. In fact, in the historic period under consideration, it may not be too much of an overstatement that court intervention in Southern corrections represents something akin to a revolution from above — that is, an effort by federal courts to modernize Southern corrections and bring them into the mainstream.[5] This greatly confounds the analysis, because it is not clear what the instrument is picking up in addition to slowed prison growth. One has to wonder whether the results would have been different had there been instrumental variables available for the large northern and western states.

A third problem of Levitt's instrument is that federal judges implementing crowding caps cannot directly affect the number of offenders sent to prison. They can urge new legislation, but short of this, the imprisonment decision remains largely with local prosecutors, judges, and juries. Remedies issued from the federal bench are primarily limited to early release or other "back end" solutions (Feeley and Rubin 1998: 387). Assuming Levitt's results are otherwise correct, this insight adds a degree of specificity to the findings: early release programs cause the crime rate to rise. Yet, we are left wondering whether front-end solutions have the same effect.[6]

To sum up, while Levitt helps us understand the consequences of court intervention under a specific set of historic conditions, we are left wondering whether the same effects occur in other contexts. Much depends on a hypothesized symmetry in causal processes between an unusual set of circumstances and a usual set of circumstances.

---

[5] On this point, see the pioneering study by Blake McKelvey (1977), and more recent contributions Paul Finkelman (1985) and Malcolm Feeley and Edward Rubin (1998: 150-158).

[6] It is easy to imagine that the effect on the crime rate of releasing one inmate early differs from that of contracting the scope of incarceration to exclude one additional inmate. For one thing, the early release may well be of an inmate with a more serious offending history. On the other hand, that inmate may no longer be criminally active (either because he has been affected by the prison experience or because he has "aged out" of offending), whereas the "new" inmate is likely to have offended in the recent past.

Perhaps these considerations suggest that we should look for better instrumental variables. Actually, we think Levitt's instrument is as close as we will come to a genuinely exogenous shifter of the prison population. Under our system of separation of powers, decisions about who goes to prison and for how long are protected from extraneous (or "non-legitimate") influences. This insulation may break down, in which case we might gain the needed instrumental variables. For example, if harsher sentences were given to minorities than to non-minorities for the same crime, percentage-minority could be used as an instrumental variable. But we are not certain of this. Judicial intervention may be the best instrumental variable we can come up with, however imperfect. In this case, what is good for the world – due process in treating like cases alike – makes it tougher for researchers to estimate their models.

C. State Level Panel Data

Perhaps no study has had a larger impact on the prison/crime issue than a paper published in the Journal of Quantitative Criminology by Thomas Marvell and Carlisle Moody (1994). The authors regressed crime rates on prison population over a 19-year period, finding that the size of the state prison population had a significant, short-term negative impact on crime. Expressed in the metric of crimes averted, "each additional state prisoner averted at least 17 index crimes on average, mostly larcenies" (1994:136). While Marvell and Moody (hereinafter "M&M") have published two follow-up papers, it is their 1994 paper that has had the largest impact and is the most formidable contribution. We focus primarily on this contribution, using it as a template for our own effort in the next section.

Like Levitt, M&M wanted to regress crime on prison. Their approach to the endogeneity issue is to examine the data for evidence of reverse causality, that is, crime driving incarceration. To do this, they employ the technique developed by Granger (1969) for analyzing causal ordering using time series data. Granger argues that one adduces evidence that variable X causes variable Y, if it can be shown that one can improve on the prediction of Y by incorporating information on the past values of X over a prediction based on past values of Y alone. In other words, if it can be shown that X helps predict Y one period ahead, beyond what can be predicted from Y's own history, then it can be said that X "Granger causes" Y.

Prior to applying the Granger causality test, M&M sought to determine whether SPP and crime are cointegrated. Given the visible upward trend in both crime and SPP, there is concern over possible spuriousness of regression results. If crime and SPP are cointegrated, they form a system in which neither variable wanders too far from the other, and the relationship between them can be separated into short-run fluctuations around a long-run equilibrium. They establish the condition of co-integration by conducting a test for a unit root in the residuals from a regression involving crime and prison population. M&M found that SPP and crime are indeed cointegrated. The Granger causality test is applied to the short-run impact of SPP on crime. For the long-run equilibrium, they do not reject the null hypothesis that causality does not run from crime rates to incarceration, but make it moot for the consideration of the short-run. Thus, they focus on the short-run relationship between changes in prison population and changes in crime rates. M&M concluded that a model regressing crime on prison populations is well specified for determining causal relationships, at least for short-run movements in the series.

There are several problems with Marvell and Moody's approach to the prison/crime relationship.[7] First, they fail to consider auxiliary variables. Granger (1969) required that any assessment of causal ordering must take place in a context which includes all relevant information. In fact, one might argue that determining Granger causality is more sensitive to specification error than OLS regression.[8] Yet Marvell and Moody consider Granger causality tests in a simple two-variable system. Thus, their conclusions about causality are drawn from a different, and less adequate, specification than the one they use to draw their conclusions about the consequences of incarceration policy. Therefore, when they reject that crime "Granger causes" SPP they have ignored the more complicated specifications including other independent variables -- determining bivariate Granger causality in a multivariate world. More properly, Granger causality analysis should include the entire information context, yielding a sort of "Granger causality net other causes."

Second, M&M's decision to punt on the long term relationship between SPP and crime is both unwise and unnecessary. It is unwise because most observers are interested in the long-run relationship between prison population and crime rates. While information on short run fluctuations has its own intrinsic value, it is the long-run impact of SPP on crime that grips us, both theoretically and for policy considerations. For example, imagine the following scenario: increasing the imprisonment rate has a short term impact on crime through the mechanism of incapacitation. But this effect dissipates because those imprisoned are "replaced" by new offenders in the medium to long run. One can have some traction in the first step up a slippery slope, but then slide down in trying to go further. Knowing only about the first step is not very useful. In sum, what is gained in M&M's methodological neatness -- in considering the prison crime relationship only in the absence of endogeneity -- comes at too high a price in theoretical and policy relevance.

Third, in our view, the Marvell and Moody model is not fully specified. The specification problem can be broken down into the inclusion of relevant control variables and the error structure used. With regard to the former, M&M used only a single set of control variables, the sizes of three age cohorts. A more complete specification would include two types of control variables: (a) states' economic and social characteristics and (b) alternative sanctions, that is, information on the number of probationers and parolees. With regard to the error structure, Marvel and Moody include a lagged dependent variable on the right-hand side of their model,

---

[7]Marvell and Moody are correct to point out that tests for Granger causality have been applied widely in economic and business studies. Some dispute its value (e.g., Maddala 1997). This issue is beyond the scope of this paper.

[8]Granger (1969: 429) is careful to point out that spurious causality cannot be assessed when relevant information is left out.

ostensibly to control for dynamic effects and autocorrelation. In our view, the problem of autocorrelation can be more effectively handled by directly including an autoregressive term into the error structure of the model.

II. EMPIRICAL RESULTS: STATE PANEL DATA

A. Data

Our analysis of state-level panel data on prison populations and crime rates covers the years 1971 through 1997, for all 50 states and the District of Columbia, unless otherwise stated. (The data sources are listed in the Appendix.) The dependent variable is the state crime rate per 100,000 population for UCR Index Crimes. This variable is logged to minimize the impact of a few large values. The key independent variable is the prison inmate population per 1000 state population, lagged by one time period. State totals include inmates under jurisdiction of both state and federal authorities.

Control variables are divided into two sets: basic and extended. By "basic" we mean the control variables used in Marvell and Moody (1994). Specifically, the variables are the percentage of a state's population in the age groups of 15 to 17, 18 to 24, and 25-34.[9] The idea here is to see if we can replicate Marvell and Moody's findings.

We include the extended control variables to account for the influences that economic and socio-demographic factors may have on crime rates. Two economic variables are included in our analysis. The first is the state unemployment rate. It is commonly argued that when unemployment increases, individuals struggling at the lower end of the socioeconomic ladder will turn to criminal activity to sustain themselves or, perhaps, be less "regulated" by the labor market. The second economic control variable is a measure of prevailing mean state wages and salaries for those with a high-school education or less, corrected for inflation.[10] Gould and

---

[9]They are constructed from U.S. Census data on the distribution by single year of age for each state. For example, the total number of persons in each state for each year of age between 15 and 17 is totaled and divided by the total state population to create the "percentage age 15-17" variable.

[10]This variable is constructed in a manner following Gould and colleagues. The yearly March CPS data from 1969-1997 were used to determine prevailing mean wages and salaries. Only males above the age of eighteen in the data with a high-school diploma or less were included. First, the income data were adjusted for inflation by using the national level Consumer Price Index (CPI-U) to 1997 dollars. Then, weighting by CPS weights, adjusted wage and salary income were regressed on dummy variables for race (white vs. nonwhite) and marital status (single vs other, married vs. other), and variables for age, education in years, and a proxy for labor force experience ((age-educ)-6)^2. These regressions were done for each year separately. The weighted mean of the residuals from this procedure are then the measure of the prevailing wage and salary in each state.

colleagues (Gould, Weinberg, and Mustard, 1998) have argued that, to predict crime rates, wages are a more important measure of labor market conditions than the unemployment rate. Wages, they argue, are a better measure of the opportunity cost of crime because the prevailing wage for unskilled workers is the true alternative to illicit gain. We find this argument sufficiently strong to include a measure of wages in addition to the more commonly used unemployment rate.

Two additional variables in the extended set of controls are the percentage of the state population who is (a) African-American and (b) reside in metropolitan areas. African-Americans are over-represented among both perpetrators and victims of crime. Urban areas have higher levels of crime, possibly due to the concentration of viable crime targets in a small area.

Finally, in some portions of the analysis we include three additional variables capturing the use of sanctions that are alternatives to incarceration. These variables are the number of persons on probation per 100,000 population, the number of persons on parole per 100,000 population, and the number of persons in jail per 100,000 population. The probation and parole data come from the Bureau of Justice Statistics. The jail data were constructed from the National Jail Census (1970, 1972, 1978, 1983, 1988, 1993) and the National Survey of Jails (1986, 1987, 1989, 1990, 1991, 1992, 1994, 1995, 1996, 1997). These data were used to compute each state's total jail population totals for the years available. For the years not available in either data set, state-specific linear interpolations were used to estimate the jail population.

Table 1 presents summary statistics for all the variables used in the model estimations below. The first column gives the mean of each variable, across all states and all years. The second column gives the standard deviation of each variable, calculated across all states and all years. The third column gives the standard deviation of state means, while the final column gives the standard deviation of yearly means. These last two columns provide some sense of the relative variability across states and across years.[11] The most striking aspect of the table is the greater variability from state-to-state than there is from year-to-year.[12]

---

[11]For virtually all variables, the state and year mean standard deviations are both smaller than the overall standard deviation; this occurs because in calculating a mean for each state or each year some variability is lost.

[12]The potential for collinearity to impact estimates was ascertained by examination of variance inflation factors (VIFs) and the condition indexes from an eigen decomposition of the matrix of independent variables. Very large VIFs, that is, in excess of 4, turned up for one of the three age demographic variables measuring the percentage of a state's population between the ages of 15-17, 18-24, and 25-34. The next largest VIF was 4.00 (for lagged prison population), but it was not associated with a large condition index. Two condition indexes exceeded 35, but again one was associated with the age demographic variables, while the second one indicated some collinearity between the proportion population black and proportion population living in metro areas. These results indicate that collinearity is not particularly problematic in our data and will not adversely affect regression results.

B. Method

The basic regression model we use is a fixed-effects panel model of the form:

$$\ln(crime\ rate_{it}) = b_0 + b_1(SPP_{i(t-1)}) + \sum b_k X_{kit} + \delta_i + \lambda_t + \varepsilon_{it} \qquad (1)$$

where $\ln(crime\ rate_{it})$ is the natural logarithm of the Crime Rate for the *i*-th state in the *t*-th year, $SPP_{i(t-1)}$ is the one-year lagged value of the prison inmate population for the *i*-th state in the *t*-th year, and δ_i and λ_t are fixed state and year effects. We examined the crime variable and prison population variable for stationarity by conducting a panel augmented Dickey-Fuller (ADF) test for each series.[13] From these tests, we feel that the crime data do not contain a unit root, but it is possible that the prison population data contain a unit root. If this is so, then there

---

[13]Testing for unit roots via the ADF is not a straightforward procedure. The underlying data-generating process may be a simple unit-root process, or a unit-root process with a non-zero mean, or even a unit-root process with a deterministic time trend. Each of these processes can be combined into a series of nested hypotheses to be tested. Additionally, there may be auto-regressive errors in the data of an unknown order, or lag length. While the ADF test allows for the specification of auto-regressive errors, the choice of the lag length is crucial; ignoring or under-guessing the actual number of lags to consider renders the unit root test standard errors incorrect. Because we do not know the proper lag-length, several lag-lengths can be used, but the results of the unit root test can change depending on the number of lags included. This means there will always be some ambiguity in drawing conclusions from unit root tests. Additionally, the series we consider are rather short compared to the long time-series typical in economics when there is usually a minimum of 100 time periods or more, making the entire ability to validly estimate the value of the root questionable. To add yet another complication, these unit root tests have low power to distinguish a unit root $\rho=1$ from a near-unit root like $\rho=0.95$. Spurious regression results from the unit root case, but not the near-unit root case. Granted, combining the information in the multiple series that form the panel data improves the ability of the test to distinguish these two cases, but this fact does not provide a full measure of comfort in the presence of the other problems of proper lag length and short series. We proceed with the test, but maintain a skeptical view of its conclusions.

ADF tests were conducted separately for each state, and the panel ADF test was constructed by combining the p-values for the ADF test from each state. As discussed in Maddala and Wu (1996), this new test statistic, the Panel ADF statistic $P_\lambda = -2 \Sigma \ln p_i$ is distributed as a chi-square statistic with df=2N, where N is the number of states. For the crime-rate data, assuming a lag length of zero (no auto-regressive errors present) leads to a rejection of the unit root null hypothesis (P_λ=148.07, df=102, p<0.01) in favor of a process with a non-zero mean. For a lag-length of one (1st order auto-regressive error present), the unit-root null hypothesis is also rejected (P_λ=277.64, df=102, p<0.01) in favor of a process that is trend-stationary. Longer lag lengths, up to eight, either lead to contradictory results, or rejection of a unit root in favor of trend-stationarity. For the prison population data, lag lengths of zero to four lead to contradictory results, while tests with lag lengths from five to eight fail to reject the unit root hypothesis.

is no need for a cointegration analysis like that used by Marvell and Moody. There is no danger of spurious regression results if crime is regressed on prison population.

We do not undertake any specific or extensive procedures to eliminate the possible simultaneity problem, save the use of the lagged value of prison population. We believe that the various approaches, such as Levitt's instrumental variable approach or Marvell and Moody's Granger test, do not satisfactorily alleviate the potential problem. Indeed, we are skeptical of the possibility that any approach can accomplish this daunting task with correlational research designs and data. Control variables are represented by X_{kit}, while ε_{it} is the residual error term. It is useful to note that in this specification, estimated coefficients of prison population and other covariates are semi-elasticities, and can be interpreted as the percent increase or decrease in crime rates due to a unit-increase of the independent variable. For some estimated models, the natural logarithm of SPP will be used, and in those cases the coefficient will be an elasticity.

The fixed effects δ_i and λ_t are indicator variables for the *i*-th state and *t*-th year. These dummy variables control for state-specific and year-specific effects that are not included in the model. The error term is assumed to be normally distributed with mean $\mu_\varepsilon = 0$ and variance σ^2_ε. In random effects models, the δ_i and λ_t are assumed to be random variables with means $\mu_\delta = 0$, $\mu_\lambda = 0$, and variances σ^2_δ and σ^2_λ. In some models, the error term is modeled as a first-order autoregressive – AR(1) – process with $\varepsilon_{it} = \rho\varepsilon_{i(t-1)} + \upsilon_{it}$ where υ_{it} is now the usual normally distributed error. Estimation proceeds via maximum likelihood.

C. Results

1. Replication and Extension of Marvell and Moody (1994)

Our initial analyses establish the sensitivity of the regression results to changes in time period covered, as well as to modest changes in model specification. These results are presented in Table 2. The first model is an approximate replication of the regression results in Marvell and Moody (1994: Table 4). The essential differences are that Marvell and Moody estimated an error correction model for the short-run relationship between inmate population and crime rates, while our model is estimating the long-run relationship. Additionally, Marvell and Moody included the lagged form of the dependent variable on the right-hand side of the model to account for dynamic effects and autocorrelation. We address autocorrelation in the next set of models.

Comparing Marvell and Moody's results with ours, both estimate a negative effect of prison population on crime rates. However, while Marvell and Moody found a highly statistically significant effect, ours is not significant. The t-ratio for our effect is t=-0.92 which has a corresponding p-value of 0.36, which is far from the standard significance level. More interesting, and more contradictory with their results, is the effect of prison population in Model 2, where we add eight years (1990-1997) to the analysis. Here, the effect is *positive* and quite significant (t=6.01). The drastic change in the effect of prison population with the addition of these years of data is surprising. In estimating a regression model, one assumes that a coefficient is constant over time. The present result suggests that a model where the effect of incarceration changes over time might be considered. Finally, in Model 2 the proportion of a state's population aged 18-24 has a positive effect on crime, and the proportion aged 25-34 have a negative effect on crime. But note that these are the variables that have collinearity issues, and significance tests of these demographic variables are suspect.

Model 3 adds the extended set of control variables to the analysis in Model 1, with the shorter time period. Note that the models are not formally comparable because some observations drop out due to missing data on unemployment in 22 states prior to 1977. What is most apparent here is that again, as in Model 1, increases in prison population are associated with lower crime rates. In Model 3, however, the effect is statistically significant. In this model, two of the age controls have non-zero effects and all of the extended set of control variables are statistically significant. The percentage of the population aged 18 to 24 is positively related to crime rates (b=2.473, t=3.60), while increases in the percentage of a state's population aged 25-34 are associated with increases in crime (1.804, t=2.38). Increases in unemployment are associated with increases in crime, as are increases in the urbanization of a state's population. These two effects follow the general expectation for them found in the criminological literature. Interestingly, increases in the prevailing wages of non college-educated persons have no effect on crime rates. This is not what we would expect if the variable taps the opportunity cost of crime. Finally, the proportion of state's population that is black is inversely related to crime rates.

The final column in Table 2 presents the same specification as in Model 3, but for the entire time period from 1972-1997. Prison population is unrelated to crime rates (t=1.41). The effects of the various control variables in Model 4 are similar to those found in Model 3, save for the significance of prevailing non college-educated male wages.

To summarize, of the four models in Table 2, prison population was non-significant twice, once it had a statistically significant positive effect, and once a statistically significant negative effect. This volatility in the estimated effect of prison on crime rates is due only to the length of the time period covered and the presence or absence of an extended set of control variables neglected by Marvell and Moody. The robustness of an estimated negative coefficient for prison in the Marvell and Moody analysis is now under some question. Our next analysis further explores this problem by changing the specification of prison population, and by incorporating autoregressive error structures in some models.

2. Investigation of Model Specification

Table 3 presents the results of various different specifications of model structure: random effects rather than fixed effects, an elasticity specification for prison population, the incorporation of a 1st order autoregressive error structure, and estimation of an error correction model. The purposes here are two-fold. First, we will be able to continue to demonstrate volatility of the estimated impact of prison population on crime. Second, we will begin to identify our preferred model specification. To accomplish the second task, we will rely on the information criteria

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) because most of the models are non-nested, abrogating the use of likelihood-ratio tests. These information criteria are penalized versions of the log-likelihood, where BIC penalizes the estimation of additional parameters more than does the AIC.[14] The smaller the value of AIC or BIC, the better fitting the model.

The first model in Table 3 is just a repeat of Model 4 from the previous table, included here to establishes a baseline for comparisons. Recall that the effect of prison population was positive and not significant. The first specification modification explored in Table 3 is presented in Model 2, where the crime rate is regressed on the lag of the natural logarithm of prison population, resulting in an elasticity form of the relation between crime and prison population. Here, the coefficient of prison population is -0.166 and highly significant (t=-9.20). The magnitude of this effect is strikingly similar to that found by Marvell and Moody (-0.159), although the model estimated here does not include an error correction term, nor any accounting of autocorrelation. Since the specification here is one of elasticities, the magnitude of the coefficient indicates a 1.66% drop in crime rates for a 10% increase in prison population. Other significant covariates in the model include the positive effects of all three age groups. Increases in unemployment seem to increase the crime rate (b=0.017, t=7.11), while increased proportion of a state's population that is African-American is significantly related to reductions in the crime rate (b=-0.043, t=-9.84). The more urbanized a state population, the higher the crime rate (b=0.006, t=5.38). Prevailing wages do not seem to affect crime rates. Model 2 is preferred to Model 1 by both the AIC and BIC criteria.

Turning to Model 3 in Table 3, we see a random effects version of Model 1 in Table 3: the state-specific and year-specific fixed effects in Model 1 are treated as random effects in Model 3. All the control variables have similar effects as in Models 1 and 2. The most notable difference between Model 3 and Model 2 is that the sign of the coefficient for prison population switches, yet the effect remains statistically significant (b=0.024, t=5.49). This implies an increase in the crime rate of 0.024% for every additional 100 persons incarcerated per 100,000 population. At the mean of SPP in these data (4950.50), a 10% increase in SPP would be 495 additional inmates, which would translate to an increase of 11.88% in the crime rate. Thus, with just the first three models in Table 3, we have the effect of prison population not significant, significantly negative, and significantly positive. Since the model specifications are not drastically different (fixed effects, random effects, monotonic transform of the prison population variable), this does not demonstrate a good deal of robustness in estimated effects of prison population. Specification 3 is not chosen by either the AIC or BIC when compared to Model 2.

We applied Durbin-Watson tests for up to 5th-order autocorrelation to the residuals from Model 1 and Model 3 in Table 3, on a state-by-state basis. Only three states failed to reject the null hypothesis of no autocorrelation for each set of residuals. This was clear evidence of at least

---

[14] The value of AIC $=-2\ell+2p$, and BIC $=-2\ell+p\ \ln(n)$, where ℓ is the value of the log-likelihood function, p is the number of parameters estimated, and n is the sample size.

a 1st order autoregressive process. For 42 of the 51 states, the Durbin-Watson test indicated the possible presence of a 2cd order autoregressive process. We then used the Smallest Canonical correlation method (SCAN) to tentatively identify the orders of any underlying ARMA process in the residuals for the two models, on a state-by-state basis. This method, which analyzes the eigenvectors of the correlation matrix for a specified ARMA process, was proposed by Tsay and Tiao (1985), and a useful description of the algorithm can be found in Box et al. (1994: 197-199). The most commonly identified ARMA structure for the 51 states was AR(1) – a 1st order autoregressive structure. Since the predominate structure identified by the SCAN approach was similar to the Durbin-Watson test results, we will incorporate a 1st order autoregressive structure in most of the subsequent analyses.

Model 4 is the first such analysis. It takes the fixed effects model and incorporates the 1st order AR(1) error structure. In doing so, there are a few changes in coefficients estimated for the control variables. The key difference is that the percentage African American in a state is no longer significant. The proportion of a state's population in urbanized areas still has a positive and significant relationship to crime rates (b=0.006, t=5.16). Of primary interest, the estimated effect of prison population remains positive, but is not significant (b=0.004, t=0.52). This model, incorporating the autoregressive error structure, is clearly preferred by both the AIC and BIC relative to any of the earlier models.

The next model in the table, Model 5, returns to the elasticity specification, but here incorporates the AR(1) error structure. In comparing the estimates in this model to those in Model 2, the size of the urban population continues to have a positive, significant effect on crime rates (b=0.006, t=5.06), while unemployment increases crime (b=0.006, t=2.55). Prison population, which had a significant negative effect in Model 2, now is not significant (b=−0.037, t=−1.59). Once again, a change in specification -- here the inclusion of an autoregressive error structure -- leads to a change in the estimated effect of prison population on crime rates. Model 5 is slightly preferred by both the AIC and BIC, relative to Model 4.

The next column in Table 3, Model 6, adds the autoregressive error structure to the random effects Model 3. In doing so, the statistically significant positive effect of prison population in Model 3 loses its significance (b=0.011, t=1.29). The control variables maintain the same pattern as in Model 3, with unemployment having a positive, significant effect on crime (b=0.005, t=2.33), as does the percentage of population in metro areas (b=0.006, t=5.16). Two of three of the included population age groups have significant associations with crime (the percent aged 15-17 having no effect and the other two positive effects). Comparing Model 6 to Model 5, the BIC selects Model 6, while the latter model is chosen by AIC. We prefer Model 6, because of the drive toward parsimony inherent in the greater penalty the BIC imposes for estimating many parameters.

While the elasticity specification in Models 2 and 3 of Table 3 comes close to the elasticity specification utilized by Marvell and Moody, it does not completely match. Having found that prison population and crime are cointegrated, they focus on estimating an error correction model for the short-run elasticity of incarceration. To match their analysis, we include an error correction form in Model 7, with fixed effects, and Model 8, with random effects. In both cases, the results are nearly identical. For the fixed effects model, the short-run elasticity of incarceration is -0.082 (t=-3.04), meaning a 1% increase in the prison population leads to a 0.08% decline in crime rates. For the random effects model, the elasticity is -0.058 (t=-2.15). These elasticities are much smaller than those reported by Marvell and Moody.[15] As to model selection, the AIC still prefers Model 5, the fixed effect model with an elasticity specification and AR(1) error structure, to any of the other models in Table 3. The BIC criterion, more heavily penalizing the use of many parameters, prefers Model 6 to either of the error correction models.

Taking an overall view of the various models and specifications included in Table 3, we can select one specification ahead of the others. Favoring the BIC, we argue that Model 6 is the preferred model from this set. This is the model that contains random effects and an autoregressive structure, with the ordinary lagged form of prison population on the right-hand side.

But there are more general statements we can add. First, the effect of prison population is *not* robust with respect to the specification of the model. The results in Table 2 indicated that depending on the time frame covered in an analysis, or on the set of control variables included in an analysis, the estimated effect of prison population was significantly positive *or* significantly negative. This pattern was once more seen throughout Table 3, where the inclusion of fixed versus random effects, autoregressive error structure or not, or using the natural logarithm of prison population, all change the nature of the estimated effect of incarceration. In some specifications, the effect is significant and negative (Model 2, Model 7, and Model 8). In one the effect is significant and positive (Model 3). In still others, prison population has no significant effect at all (Model 1, Model 4, Model 5, and Model 6). This behavior is in strong contrast to the robustness of the estimates of the control variables, most vividly seen in Table 3. The magnitude and significance of the effects of control variables rarely changed from model to model. That the effect of prison population *did* change from model to model compels researchers to rethink the specification of their models.

Second, where there is evidence of autocorrelation, as often occurs in time-series data, it should be taken into account in the model specification. We encourage actually modeling an autoregressive error structure rather than merely including a lagged form of the response variable on the right-hand side, which borders on an ad hoc solution to the problem. When the error term is specified, estimates of its effects are more likely to be reliable and problems with estimation will be more transparent.

---

[15]For comparative purposes we ran the same two error correction models on the data for the same time period, including only the age demographic set of controls as had Marvell and Moody. We included an AR(1) effect to match up the inclusion of the lagged value of crime they used to capture any autocorrelation. This leads to an estimated elasticity of 0.003 for the fixed effects model, and -0.005 for the random effects model. Neither effect was significant.

Third, there may be a slight advantage in fit by using the elasticity specification. The advantage is very slight indeed. The difference in the value of the log likelihood, as well as the AIC and BIC, between Model 1 and Model 2 is 61.5, or only a 3 or 4 percent difference in the fit statistics between the models. Once the autoregressive error structure is included, the improvement shrinks to a mere 2.7 (about one-half of 1 percent) in the value of the log likelihood, AIC, and BIC (Compare Model 4 with Model 5).

3. The Role of Alternative Sanctions

Turning to Table 4, we continue our analysis in a novel way, by considering alternative sanctions available to criminal justice authorities. Rather than incarcerating a guilty individual, courts can place the offender on probation. Rather than keeping a prison inmate for the full length of the term, parole boards can put them back on the street. Judges can sentence offenders such that they are incarcerated in jails rather than in prisons. These options are included as controls in this part of the analysis. In Model 1 and Model 3 of Table 4, a fixed effects model is estimated, with the difference between the two being the addition of an AR(1) error structure to the latter model. Similarly, Model 2 and Model 4 are random effects models, the latter with an AR(1) error structure. Note also that probation and parole data are not available for years prior to 1977, nor jail data for all 51 "states," resulting in fewer observations.[16]

Most striking in Table 4 is the absence of any effect of prison population. It is not significant in any of the four models presented. Similarly, the alternative sanctions of probation, parole, and jail do not seem to have an effect on crime rates, save for the significant positive coefficient of probation populations in the first two models and parole in the second model. But those effects disappear once the autoregressive error structure is added in Models 3 and 4. Looking at the control variables, a now common pattern is readily seen. The urbanization of a state's population continues to have a significantly positive effect on crime rates, and unemployment has a weakly significant positive effect. The prevailing wage variables are significant in the first two models, but those effects fade away in the final two models. This continuing stability of the control variables, unlike the sanction variables, suggests that the instability is not a statistical artifact.

Why do virtually none of the sanctions have any effect? One technical possibility is collinearity, but as discussed earlier, there is no evidence of that problem in these data. Another possibility is that the various sanctions have specialized effects on specific types of offenses, and those effects wash out when the response variable is all index offenses. Thinking specifically about probation, that particular variable may be cutting too wide a swath – many people receive probation for misdemeanors. A sharper measure of probation (distinguishing felony probation from misdemeanor probation) might yield different results. For a similar reason, jail population may not have any effect on the rate of index crimes, since jail terms are generally given for serious misdemeanors or for felons that receive the lowest of sentences.

---

[16] The excluded states are Alaska, Connecticut, Delaware, Hawaii, Rhode Island, and Vermont.

4. Extensions to the Basic Model

In Table 5, we extend our exploration of prison population by three additional specifications, first with fixed effects and second with random effects, both with an AR(1) error structure. The first additional specification, presented as Models 1 and 2, attempts to discern any threshold effects by the inclusion of a quadratic form of prison population. A threshold effect would be observed if the signs on the linear and quadratic forms of prison population were the same. If both were negative, this would indicate that as the number of inmates grows ever larger, at some point the crime-reduction effect of incarceration "takes off" and becomes stronger. If both coefficients are positive, then the same interpretation would hold, but the effect would be crime-facilitating. As can be seen in the first two columns of the table, the sign of the linear term is negative and the sign of the quadratic term is positive for both the fixed and random effects model. This would be indicative of an U-shaped relation; increases in prison population decrease the crime rate, but as the number of inmates continued to grow, the effect would weaken, and at some threshold reverse. An F-test on the joint significance of the linear and quadratic term is not significant ($F_{2,803}$=1.46, p<0.23 for the fixed effects model; $F_{2,847}$=0.91, p<0.40 for the random effects model).

In Model 3 and Model 4, state-specific time-trends are included to capture any effects that are difficult or impossible to measure, but that vary from state-to-state. One example might be a general trend toward or away from incarceration for certain classes of offenses, such as might happen due to judges responding to a public outcry. Perhaps a state has undergone an investment in improved technology, which can directly affect the crime rate. Even something so simple as an improvement (or destruction) of police benefit packages might indirectly impact the arrest rates and thus, indirectly, crime rates. Because the time-trends are state-specific, there are actually 44 such effects in the models, and thus not included in the table. An F-test for the joint significance of these effects is significant for both models ($F_{44,760}$=1.96, p<0.0003 for Model 3; $F_{45,804}$=1.87, p<0.0006 for Model 4). Does the inclusion of the state-specific time-trends impact the estimated coefficient of prison population? Apparently not, as the coefficient is negative, but not significant in both models.

Finally, in Models 5 and 6 of Table 5, we explore yet another interesting specification. Traditionally, the effect of prison population is estimated as a fixed effect, constant for all states and all time periods. In these last two models, we allow the effect of prison population to vary over time.[17] In the fixed effects Model 5, this is accomplished by including an interaction of the year dummies with the prison population variable. Thus, the effect of incarceration is the same for all states within a single year, but it changes from year-to-year. With this specification, an F-test can be conducted on the joint significance of the varying coefficient. This specification turns out to be statistically significant ($F_{19,785}$=6.14, p<0.0001). The value of the coefficient for each year is presented in Table 6. The overall pattern is that the effect of prison population grew

---

[17]Note that Marvell and Moody did identify a change in magnitude in the effect of state prison population, with the change pegged to 1975, and increasing in the later years.

somewhat through the late 1970s and first half of the 1980s, then began a slow decline through the rest of the decade and on into the 1990s, until 1997. In 1997, an abrupt increase in the magnitude of the effect of prison population is observed.

For the random effects version, Model 6 in Table 5, the time-varying component of the prison population effect is viewed as a random variable that fluctuates around an overall time-invariant fixed coefficient. The value of the time-invariant mean coefficient of prison population is -0.010, and is not significant (t=-1.41). To better compare the time-varying coefficients with the fixed effects model, the time-invariant effect of prison population in Model 6 is added to the time-varying fluctuations and presented in the third column of Table 6. The pattern of the total effects from Model 6, although only correlating r=0.48 with the varying coefficients from Model 5, do seem to follow a similar pattern as the varying effects in Model 5. This similarity is most strongly seen in the 1990s, and much less so for the earlier years. The strong negative effect in 1997 from Model 5 is echoed in the strong negative effect in 1997 from Model 6.

III. EMPIRICAL RESULTS: FLOW DATA FROM THE NATIONAL CORRECTIONS REPORTING PROGRAM

A. Data

In this section, we continue with a state-level panel analysis, but exploit data from the National Corrections Reporting Program (NCRP). The NCRP collects data on every admission and release from prison for those states that report. The program began in 1983, and the most recently available data are from 1996. For each admission, demographic information (age, sex, race, and ethnicity), prior incarceration history (prior jail time, prior prison time), and current offense and sentence information are recorded.

There are three advantages to the NCRP data. First, they allow us to analyze the flow of inmates into prison instead of (or in addition to) the stock measure of prison population used in the previous section. This may be enlightening, because we will now use a flow measure (admissions) to explain a flow measure (crime rates). Whether flow matters more than stock in crime reduction is an empirical question. For example, if prison reduces crime through the mechanism of deterrence, then the key question is this: does the potential offender calculate (however crudely) the cost of crime based on the number of offenders who went to prison last year or from the total number of prisoners? In fact, if a potential offender was rationally calculating and omniscient, flow would be more important to him or her than stock. The best predictor of next year's incarceration rate is last year's incarceration rate. On the other hand, an offender from a high incarceration-rate community might calculate the costs of crime based upon the number of community members who have been sent to prison over the years. Critics of high incarceration rates often point out that such policies can literally remove a large proportion of young males from a community; this fact may not be missed by those who live in those communities. If the latter is the case, the stock may have a larger impact on crime than the flow.

Second, the NCRP data provide information on each prisoner's incarcerating offense. We are particularly interested in separating drug offenders from other felony admissions, and then determining whether imprisoning drug offenders has the same effect on crime as imprisoning non-drug offenders. Third, the NCRP data can be used to construct a measure of sentencing practices. This, in turn, allows us to determine the effect of incarceration net of the severity of sentencing. In other words, we will be able to separate out the impact of an increase in prison population arising from the widening of the net from an increase due to longer sentences to the same types of offenders as incarcerated previously.

Each year, approximately 35 states participate in the NCRP by reporting data on each admission to and release from prison and parole. Some states have short or intermittent association with the program. Only seventeen states have been part of the program since its inception. Nine states have never supplied data to the NCRP. A full breakdown of which states have reported data for which years is presented in Figure 1. Our subsequent analyses use as many state/year observations as available.

Aggregate state-level counts of the number of admissions are created from the master data files by summing the number of cases (admissions) by reporting state and reporting year. Likewise, data on the most serious offense committed – identified as the offense with the longest sentence – are used to classify every prison admission by offense type (in particular, drug vs. property or violent). We point out that the categorization of offense types is not obvious or simple. Our definition of a drug offender is one whose most serious offense was a drug offense. However, alternative definitions of "drug offender" could include everyone admitted to prison with any drug offense or could distinguish between possession and trafficking.

Given our definition, the number of admissions by offense type for every state-year is then determined in the same manner as total admissions. There is some loss in state-year level data at this point, since offense data are missing for some states in some years. For instance, Arkansas in 1990 and 1991 has no offense information for 32% and 43% of admissions, respectively. When the analysis breaks out admissions by offense type, we drop observations with a high proportion of missing data.

Sentencing practices are of interest because prison populations can expand in two ways. A larger proportion of convicted offenders may be sent to prison or the sentences given out may be increased in length. In fact, it may be that sentencing practices are the driving force explaining variation in crime rates. Under this perspective, incarceration may be an imperfect and incomplete proxy for sentencing; not accounting for sentencing could lead to spurious effects of incarceration. Sentencing -- as a manipulable policy alternative akin to probation -- is of interest on its own merits and not merely as a control variable. However, in the current research we will be primarily interested in sentencing as a control.

B. Method

Our initial strategy is to replicate the results presented in Table 3, but substituting the flow measure of new admissions in place of the stock measure of state prison population. For the first analysis, total admissions per 1000 population will be included in a basic flow-flow analysis. This analysis will highlight any differences in the use of a flow vs. stock measure of incarceration.

The next analysis will separate drug and non-drug offense admissions. This will allow us to conduct the first reported study of the separate effects of incarceration by type of offense committed. The same model specifications from the previous analysis will again be used.

The final analysis will highlight the preferred models from the previous analysis, but estimated separately for violent and property crime rates. We might expect the incarceration of drug offenders to have a different impact on violent than on property crimes. To simplify the presentation, only the best fitting models from the prior set of model specifications will be used.

Constructing a measure of sentencing severity is far from a simple task. Indeed, a decision must be made whether one is interested in sentencing *policies* or sentencing *practices*. If interested in sentencing policies, one would turn to legal declarations of minimum and/or maximum sentences for different offenses and felony degrees, as well as the presence and breadth of state mandated determinative sentencing. If interest lies in sentencing practices, then one turns to actual sentences given to offenders when courts have flexibility to decide sentence length.

We believe that practices, including plea bargaining, vary much more across states and over time than do policies. States do not engage in wholesale revision of policies very often and, even when they do, the changes tend to be directed as much toward classification and reclassification of offenses as in changing the sentencing limits embedded in minimum/maximum sentences. On the other hand, sentencing practices can vary widely within the policy-mandated limits; a state attorney general may lead an attack on certain offenses by advising judges to give sentences at the upper limit of the prescribed ranges, for example. Likewise, certain offenses can be informally made less serious by suggesting sentences toward the lower limit of policy prescriptions.

The NCRP allows for the constructing of a measure of sentencing practices for multiple states over multiple years. Sentencing information is provided for prison admissions and for prison releases, including the minimum prison term to be served, the total maximum sentence length for all offenses, and the maximum sentence for the most serious offense. For prison releases, actual time served is provided.

While it may seem prudent to use actual time served as a measure of sentencing practices, there are problems with such an approach. First, given the fourteen-year history of the NCRP, reliance on prison releases means that more serious offenses that would garner long sentences will tend to be under reported as inmates with long sentences are less likely to be released during the 1983-1996 period covered by the NCRP. Secondly, sentences and time served by prison releases in any year do not actually measure sentencing practices during that year. They are releases after all and not admissions. Third, using the release data to construct retrospective sentencing data for the year that the prison release was initially admitted will lead to a hit-and-miss picture of sentencing during admission years, with an incomplete non-representative set of sentences for the admission year (those serving the longest terms will tend to be under represented, as those inmates are less likely to be released, which leads to "survivorship" bias).

Therefore, we use sentencing information from prison admissions to construct our measure of sentencing practices. We do this in a manner similar to the construction of our prevailing wage measure. Separately by year, we regress a person's maximum sentence for the most serious offense on age, education, prior prison time served, and offense. We include only males in the estimation. The residuals from this regression give variation in sentences which are net of age, education, offense type, and prior time served. A negative (positive) residual indicates a sentence given that is lower (higher) than that given on average to a person with the same characteristics of age, education, offense, and prior time served. Next, we group these residuals by state to determine the state mean. A state with a negative mean residual is one which tends to give shorter sentences in practice than other states for persons with the same offense and prior history. A state with a positive residual tends to give longer sentences in practice than other states.

C. Results

We first replicate the state-level analysis of the earlier section, but with a flow measure of state prison population in place of the stock measure used in the earlier analysis. While these two measures are highly correlated in levels (r=0.879), there are differences in their substantive meaning. The standard stock measure of state prison population is, in one sense, a weighted sum of prior yearly admissions into prison, where the weight is the average sentence length. Crime, our response variable of interest, is a flow measure -- the rate at which crimes are committed *for a single time interval*, one year. Because of this, we are interested in the influence of a prison population flow measure for a single year, admissions to prison. It is this substantive distinction which leads us to explore the state prison admissions flow measure in these analyses. The issue boils down to an empirical, rather than a theoretical, matter: does flow or stock have a larger effect on crime, and in which direction?

1. Replication of Results using New Admissions Measure

The replication begins with a set of analyses parallel to those reported in Table 3. Note that since the prison admissions data come from the NCRP the time frame of the analysis begins in 1984 rather than in 1972 as in the earlier analysis. These analyses include a number of specifications which are presented in Table 7. The first model is a simple fixed effects model. The effect of prison admissions is not significant, while the relevant control variables are generally in the expected direction. Increases in the unemployment rate and the proportion of a state population living in metro areas are associated with increases in the FBI index crime rate. Increases in the prevailing wage for non college-educated males are negatively related to the crime rate. The state percent black population also has a negative relationship to the crime rate.

Model 2, in the next column of Table 7, estimates a model with an elasticity specification for prison admissions by using the natural logarithm of admissions. The results are identical to the non-elasticity specification in Model 1; no effect of prison admissions on the crime rate, while increases in the unemployment rate and the proportion living in metro areas are associated with increases in the crime rate. The prevailing wages for non-college males and the proportion of population black lead to declines in the crime rate.

Model 3 reverts back to the semi-elasticity specification of the fixed effects Model 1, but substitutes random effects for the fixed effects. The results are strikingly similar to the fixed effects model, as well as to the elasticity specification Model 2. The only change is that percent black is no longer significant. There is no discernable effect of prison admissions.

In Models 4 and 5, the fixed effects specification is augmented with a 1st order autoregressive error structure. Additionally, an elasticity specification for prison admissions is included in Model 5. In Model 6, the AR(1) error structure augments a random effects model with no elasticity form for prison admissions. For all three models, prison admissions -- entered both in levels (Models 4 & 6) and elasticity (Model 5) forms -- does not reach standard levels of significance. The effects of the various control variables have not changed substantively from the earlier models: increases in unemployment and metro population lead to increases in crime rates, while increases in wages for non-college males lead to reductions in the crime rate.

The final two columns utilize an error correction model specification: Model 7 including fixed effects and Model 8 with random effects. The short-run effect of prison admissions is not significant in either model. Somewhat unexpectedly, increases in unemployment are associated with declines in crime rates. The proportion metro population still leads to higher crime rates.

Choosing the best models from the set of eight is not a straightforward proposition. The likelihood ratio tests cannot generally be employed because most of the models are not nested. As in the earlier analyses, we can use the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to compare models. The larger the negative value of AIC or BIC, the better the model. Recall that the BIC imposes a stronger penalty for fitting lots of parameters. Models 4 and 5 are the best models in Table 7 using these two information criteria. Note that in both of these models -- fixed effects models incorporating AR(1) error structures and an elasticity form in Model 5 -- the effect of prison admissions is not statistically significant.

The overall impression from this set of analyses is that although the stock and flow measures of prison inmates are highly correlated, they clearly are not that similar when used to model crime rates. For the stock measure, state prison population was significant in several of the models (albeit with a direction that bounced around) whereas the flow measure of prison admissions never achieved statistical significance.

2. Decomposing Admissions by Offense Type

One of the advantages of the NCRP data is the ability to separate out drug offender admissions from other admissions. We redo the analysis just discussed using this disaggregated measure of flow in hopes of better specifying our models. Certainly, recent debates in criminology have frequently speculated on the role of the enhanced "war on drugs" of the past decade on crime rates. Briefly, many and perhaps most criminologists argue that imprisoning drug offenders will have little effect on the crime rate. For one thing, imprisoning a drug offender may simply open a market niche for a new drug offender. On the other hand, some practitioners report an observed link between crackdowns on drug offenders and reduced crime. In recent Congressional testimony, New York City mayor Rudolph Giuliani asserts that "in areas where our anti-drug initiative zones have been in effect, we've seen the crime-declines outpace citywide decline" (1999). Separating out drug offender admissions from other admissions allows us to address the debate from a broad empirical direction. The results are presented in Table 8.

Models 1 and 2 are fixed effects models, but Model 2 uses an elasticity form for both drug offender admissions and non-drug offender admissions. Model 3 is a random effects model. The effect of prison admissions is consistent across all three specifications. The effect of non-drug admissions is to *increase* the crime rate, while increases in drug offender admissions act to *reduce* the crime rate. In Model 1, an increase of one non-drug admission per 1000 population results in a 0.047% increase in the crime rate. At the mean of non-drug admissions per 1000 population (1.11), a 10% increase would be 0.11 non-drug admissions. Thus, a 10% increase in non-drug admissions leads to a 0.052% increase in the crime rate. For Model 2, a 10% increase in non-drug admissions produces a 0.56% increase in crime rates, and Model 3 a 10% increase in non-drug admissions per 1000 population (at the mean) leads to a 0.085% increase in crime rates. All effects are significant at the 0.01 level. For drug offenders, an increase of 10% in drug offender admissions per 1000 population (at the mean) leads to a decline in crime rates of 0.004% and 0.003% in Models 1 and 3, respectively. For Model 2, a 10% increase in drug offenders admitted to prison produces a 0.50% decline in crime rates.[18] In short, these *initial* results are more consistent with the position that imprisoning drug offenders has an impact on the crime rate (as argued by Giuliani) rather than the position that drug offenders are merely replaced when imprisoned. Yet additional analyses should give pause.

The next three models, Models 4, 5, and 6, add a 1st order autoregressive error to Models 1, 2, and 3, respectively. Noting the change in AIC and BIC by adding the AR(1) error, it seems to be important to have the error structure so designated. In moving from Model 2 to Model 5 (the elasticity specification), the significant effects of non-drug and drug offender admissions disappear. More disturbingly, adding an AR(1) term to the fixed effects (see Model 4) and

---

[18]As with the models using stock measures, all three models show that increases in unemployment and percent living in metro areas increase the crime rate, while increases in percent black (in Models 1 and 2) and prevailing wage for non-college educated males decrease the crime rate.

random effects models (see Model 6) leads to a loss of significance for non-drug admissions, while the effect of drug offender admissions flips sign. What were significant negative effects in Models 1 and 3 are now significant positive effects.

The final two columns show the results for the error correction form, with fixed effects (Model 7) and random effects (Model 8). There is no discernable short-run effect of non-drug offender admissions (the coefficient of log(admissions/1000) in the two models), while only the effect of drug offender admissions in Model 8 is significant (and negative). The effects of the control variables were remarkably stable and unsurprising over Models 4 through 8. It is only the effect of non-drug and drug offender admissions that changes from specification to specification. We seem to be seeing the same sort of effect instability as we did with the stock measure analysis in Table 3, although for the models selected as best by AIC and BIC (Models 4 and 5), the effect of drug offender admissions is significantly positive in Model 4 and not significant, but positive, in Model 5.

One possible confounding element in the analysis is sentencing practice. Our sentencing measure is an attempt to quantify sentencing practice -- the actual relative length of sentences meted out, on average, by judges within a state. Our next analysis replicates the analyses in Table 8 that were just discussed, but we add our measure for sentencing practices.

The most striking result that can be observed in Table 9 is that our measure of sentencing practices never reaches statistical significance, no matter the specification of the overall model. In a nearly similar manner, the effect of either non-drug admissions or drug offender admissions is generally not significant across the eight models. The two exceptions are Models 2 and 7.

For Model 2, a fixed effects model with an elasticity specification for non-drug admissions and for drug offender admissions, the coefficient for log (non-drug admissions) is negative and statistically significant at the 0.01 level. In this instance, the effect of a 1% increase in non-drug admissions reduces the index crime rate by 0.08%. For Model 7, there is a statistically significant negative effect of drug admissions in the short-run. For a short-run increase of 1% in drug admissions, there is a short-run decline in the index crime rate of 0.046%.

As to the effect of the various control variables, a fairly familiar picture emerges. The proportion of a state's population residing in metro areas has a significant positive effect on crime rates in all but one model specification. Increases in unemployment produce statistically significant increases in the index crime rate in Model 4 and Model 5, and the effect is significantly negative in Model 8. The measures of prevailing wages for non-college educated males have no effect in any of the models.

In a final effort to discern the impact of non-drug and drug offender admissions, we analyze the effects of such admissions on violent and property crime rates separately. One possibility is that the effects of non-drug and drug offenders have been masked by consideration of overall index crime rates. More specifically, we might speculate that drug offender admissions would have more of impact on property crime rates more than on violent crime rates.

For this disaggregated analysis, we report estimates only for the three best fitting models determined by a consideration of AIC and BIC, and, to a lesser extent, the value of the log-likelihood for the models in Table 8.

The selected models are Model 4 (fixed effects w/AR[1]), Model 5 (fixed effects w/AR[1] and elasticity form of admissions), and Model 7 (fixed effects, error correction form). Given the poor showing of the sentencing practice variable in Table 9, it is not included in the disaggregated analyses. The results are presented in Table 10.

For violent crime rates, neither the non-drug admissions, nor the drug offender admissions have any statistically significant effect. Increases in unemployment have either no effect on violent crime rates (Models 1 & 2) or a negative effect (Model 3). This last effect may seem unexpected, but note that the error-correction form in Model 3 fits much worse relative to the first two models. The proportion of a state's population that is black is associated with an increase in violent crime rates, but only in the first two models. Finally, increases in metro population increase violent crime rates (all models), and prevailing wages for non-college educated males have no effect.

Turning to property crime rates, we find a significant positive effect of drug offender admissions for Model 1. An increase of 10% in drug offender admissions to prison per 1000 population (at the mean) leads to an increase of 0.002% in the property crime rate. This positive effect might be interpreted as consistent with the position that imprisonment of drug offenders makes crime worse by damaging communities and failing to address the underlying drug problems which, in turn, provide motivation for crime. This positive effect, however, is not observed in the other two models, where there is no discernable effect on property crime rates. Admissions of non-drug offenders to prison has no effect in any of the three models for property crime.

Control variable effects on property crime are different from what we have already seen. Increases in the unemployment rate lead to increases in property crimes, as expected (Models 1 and 2 only), as do increases in the concentration of population into metro areas (all three models). However, the effect of prevailing wages for non-college educated males is significant and negative in the first two models, and significant positive in the third model. Once again, though, the rather poor fit of the error correction model (Model 3) relative to Models 1 and 2 (in terms of AIC and BIC) cautions us from making too much of the positive coefficient in Model 3.

A review of the admission flow models estimated in Tables 7 through 10 does not lend itself to any ready generalizations. The basic flow models in Table 7 indicate no effect of increasing prison admissions on crime rates. A separation of admissions into non-drug and drug offender admissions in Table 8 points to a mixed bag of effects, similar to that observed in the prison stock analysis of the earlier section: drug offender admissions sometimes have significant positive effects on crime rates and sometimes have significant negative effects on crime rates. Model specification would seem to be an important component to the kind of results observed. Finally, separation of crime rates into violent and property crime rates does not clarify the issue so much. The only discernable effect in Table 10 is a positive effect of drug offender admissions on property crime rates in a fixed effects model.

We have presented a large number of analyses of the crime-prison relationship. The first set of analyses established that the most widely cited estimates of this relationship are extremely sensitive to very modest changes in the model, including adding reasonable control variables and more years of data. Our second set of analyses was a more technical investigation of various error structures. From this, we concluded that modeling the autocorrelation in the data is important. At the same time, the lack of robustness in the crime-prison relationship to the various model specifications reinforced our earlier conclusion that researchers are not currently in a position to have confidence they have isolated the true effect of prison on crime. Other analyses considered the possibility that the flow of admissions (rather than the stock of offenders) drives changes in crime rates. Finally, we considered the possibility that different "types" of offenders might have different effects on the crime rate. This latter line of inquiry has the potential to bridge the gap between aggregated studies of incarceration of the type emphasized in this paper and the survey-based approaches to the same broad question of the effect of prison on crime.

IV. CONCLUSIONS

Many now argue that the United States relies excessively on prisons. Others are heartened by the recent decline in the crime rate, and want to credit prison expansion for at least some of this trend. In our view, both positions are reasonable, and unproven. One purpose of this report has been to shake the reader's confidence in existing estimates. We have not provided a "new and improved" estimate, though we think that this is entirely possible and is the goal of our future work.

We lack confidence in existing efforts for several reasons. The key one is the lack of robustness of the regression results in work on this subject widely cited in research and policy circles. If indeed there were a deep, causal association between prison population and crime rates, minor changes in specification or the years covered would alter only slightly the results. In conducting our own analyses, we did not find this. The significance levels changed and even the signs reversed depending upon time period covered, control variables included, and the estimating techniques used. We have no way to explain these perturbations, and the only reasonable statement at this stage is that we are not sure what is going on.

Finally, it is possible that other issues raised here, such as the use of other correctional sanctions and the possibility of nonlinear impacts, matter a lot. Additionally, a different picture sometimes seems to emerge if one uses prison admission flows rather than prison inmate stock. One issue, with clear policy implications, is whether the imprisonment of a large number of drug offenders has any impact on the crime rate. There is some evidence that the effect of non-drug admissions is not the same as that of drug offender admissions. The difficulty in making a definitive conclusion comes from the inability of statistical devices to choose among different model specifications and the instability in estimated effects across those different model specifications. Future work must begin to pay attention to the choice of model specification and find justification for chosen specifications in theoretical considerations.

REFERENCES

American Legislative Exchange Council [ALEC]. 1994. Report Card on Crime and Punishment. Washington, D.C. ALEC Foundation.

Austin, James and John Irwin. 1993. Does Imprisonment Reduce Crime? A Critique of "Voodoo" Criminology. San Francisco: National Council on Crime and Delinquency.

Becker, Gary. 1968. "Crime and Punishment: An Economic Approach." Journal of Political Economy. 76:169-217.

Box, G.E.P., Jenkins, G.M., and G.C. Reinsel. 1994. Times Series Analysis: Forecasting and Control (3ed.), Englewood Cliffs, NJ: Prentice Hall.

Bureau of Justice Administration. 1998. 1996 National Survey of State Sentencing Structures. Washington, D.C.: BJS.

Chaiken, Jan M. and Marcia R. Chaiken. 1982. Varieties of Criminal Behavior. Santa Monica: Rand.

Cook, Thomas D., and Donald T. Campbell. 1979. Quasi-Experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton Mifflin Company.

Clear, Todd R. 1996. "The Unintended Consequences of Incarceration." Paper presented to the National Institute of Justice Workshop on Corrections Research, February 14-15, 1996, Washington, D.C.

DiIulio, John J. Jr., and Donald F. Kettl. 1995. Fine Print: the Contract with America, Devolution, and the Administrative Realities of American Federalism. Washington, D.C.: Center for Public Management, The Brookings Institution.

DiIulio, John and Anne Piehl. 1991. "Does Prison Pay? The Stormy National Debate over the Cost-Effectiveness of Imprisonment." The Brookings Review (fall), 28-35.

Donziger, Steven A.1996. The Real War on Crime : The Report of the National Criminal Justice Commission. New York : HarperPerennial,

Ekland-Olson, Sheldon, William R. Kelley, Hee-Jong Joo, Jeffrey Olbrich, and Michael Eisenberg. 1993. Justice Under Pressure: A Comparison of Recidivism Patterns Among Four Successive Parolee Cohorts. New York: Springer-Verlag.

Fajnzylber, Pablo, Daniel Lederman, and Norman Loayza. 1998. Determinants of Crime Rates in Latin American and the World: An Empirical Assessment. Washington, D.C.: The World Bank.

Feeley, Malcolm M. and Edward L. Rubin. 1998. Judicial Policy Making and the Modern State: How the Courts Reformed America's Prisons. Cambridge: Cambridge University Press.

Finkelman, Paul. 1985. "Exploring Southern Legal History." North Carolina Law Review. 64:77-120.

Freeman, Richard B. 1995. "The Labor Market." In James Q. Wilson and Joan Petersilia (eds.) Crime. San Francisco: Institute for Contemporary Studies.

_______________. 1996. "Why Do so Many Young American Men Commit Crimes and What Might We Do About It? Journal of Economic Perspectives, 10:25-42.

Giuliani, Rudolph W. 1999. Testimony before U.S. House of Representatives Committee on Government Reform and Oversight. "Fighting Crime at the Local Level." March 3.

Gould, Eric D., Bruce A. Weinberg, and David Mustard. 1998. "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1995." Unpublished paper, Hebrew University, August.

Granger, C.W.J. 1969. "Investing Causal Relations by Econometric Models and Cross-Spectral Methods." Econometrica. 37: 424-438.

Greenwood, Peter W. 1987. Selective Incapacitation. Santa Monica: Rand.

Hendry, David F. 1995. Dynamic Econometrics. Oxford: Oxford University Press.

Joo, Hee-Hang, Sheldon Ekland-Olson, and William Kelley. 1995. "Recidivism Among Paroled Property Offenders Released During a Period of Prison Reform." Criminology 33:389-410

Kahan, Dan M. 1997. "Between Economics and Sociology: The New Path of Deterrence." Michigan Law Review. 95:2476-2497.

Katyal, Neal K.1997. "Deterrence's Difficulty." Michigan Law Review. 95: 2385-2497

Levin, A. and C.F.Lin. 1992. "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties." Discussion paper #92-93, Department of Economics, University of California at San Diego.

_________________. 1993. "Unit Root Tests in Panel Data: New Results." Discussion paper #93-56, Department of Economics, University of California at San Diego.

Levitt, Steven D. 1996. "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation." Quarterly Journal of Economics. 111:319-52.

Maddala, G. S. and In-Moo Kim. 1999. Unit Roots, Cointegration, and Structural Change. New York: Cambridge University Press.

Maddala, G.S., Robert P. Trost, and Frederick Joutz. "Estimation of Short-Run and Long-Run Elasticities of Energy Demand from Panel Data Using Shrinkage Estimators," Journal of Business and Economic Statistics, 15: 90-100.

Maddala, G.S., and Shaowen Wu. 1996. "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test: Evidence From Simulations and the Bootstrap." Paper presented at the EML/NSF Symposium on the Bootstrap. Berkeley, July 30-Aug 6, 1996. http://emlab.berkeley.edu/eml/nsf96/nsf96.html

Marvell, Thomas B. and Carlisle E. Moody. 1994. "Prison Population and Crime Reduction." Journal of Quantitative Criminology. 10:109-39.

____________________. 1997. "The Impact of Prison Growth on Homicide." Homicide Studies. 1:205-33.

____________________. 1998. "The Impact of Out-of-State Prison Population on State Homicide Rates: Displacement and Free-Rider Effects." Criminology. 36:513-535.

Mauer, Mark..1999. Race to Incarcerate. New York: The New Press.

McKelvey, Blake. 1977. American Prisons : A History of Good Intentions. Montclair, N.J. : P. Smith

McCoskey, Suzanne and Chihwa Kao. 1998. "A Residual Based Test of the Null of Cointegration in Panel Data," Econometric Reviews, 17: 57-84.

Meares, Tracey L. 1998. "Social Organization and Drug Law Enforcement." American Criminal Law Review. 35:191-227.

Moore, Mark H. and Anne Morrison Piehl. 1998, "Reckoning the Value of Prisons: The Strengths and Limitations of Benefit/Cost Analyses of the Crime-Reduction Effects of Imprisonment," Unpublished paper, Harvard University, September.

Nagin, Daniel and Joel Waldfogel. 1995. "The Effects of Criminality and Conviction on the Labor Market Status of Young British Offenders."International Review of Law and Economics 15:109-126

Piehl, Anne M. and John DiIulio. 1995. "Does Prison Pay? Revisited." The Brookings Review, (winter), 21-25.

Piehl, Anne M., Bert Useem, and John J. DiIulio. Jr., September, 1999. Right-Sizing Justice: A Cost-Benefit Analysis of Imprisonment in Three States. A Report of the Manhattan Institute for Policy Research Center for Civic Innovation.

Sampson, Robert and John Laub. 1993. Crime in the Making: Pathways and Turning Points through Life. Cambridge: Harvard University Press.

Skolnick, Jerome H. 1995. "What Not To Do About Crime -- The American Society of Criminology Presidential Address." Criminology. 33:1-15.

Spelman, William. 1994. Criminal Incapacitation. New York: Plenum.

Tsay, R.S., and G.C. Tiao. 1985. "Use of Canonical Analysis in Time Series Model Identification," Biometrika, 72: 299-315.

Waldfogel, Joel. 1994. "Does Conviction Have a Persistent Effect on Income and Employment? International Review of Law and Economics. 14:103-119.

Western, Bruce and Katherine Becket. 1999. "How Unregulated Is the U.S. Labor Market? The Penal System as a Labor Market Institution." American Journal of Sociology. 104:1030-60.

Witt, Robert and Ann D. Witte. 1999. "Crime, Imprisonment, and Female Labor Force Participation: A Time-Series Approach." National Bureau of Economic Research. Working paper 6786. http://www.nber.org/papers/w6786.

Zedlewski, Edwin. 1987. Making Confinement Decisions. NIJ Research in Brief. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics, USGP0.

Zimring, Franklin and Gordon Hawkins. 1997. Crime Is Not the Problem: Lethal Violence in America. New York: Oxford University Press.

Zimring, Franklin and Gordon Hawkins. 1988. "The New Mathematics of Imprisonment." Crime and Delinquency 34:425-436.

Figure 1. List of years with usable data from NCRP for various states.

Alabama	1983-1996
Alaska	1994
California	1983-1996
Colorado	1983-1996
Delaware	1983-1988
Dist of Columbia	1985-1989, 1991-1994
Georgia	1983-1984, 1987, 1989-1996
Hawaii	1984-1989, 1991-1996
Illinois	1983-1996
Iowa	1987-1996
Kentucky	1983-1996
Louisiana	1992-1996
Maine	1992-1996
Maryland	1983-1996
Massachusetts	1983-1988, 1992-1994
Michigan	1983-1996
Minnesota	1993-1996
Mississippi	1984-1989, 1991-1996
Missouri	1984-1996
Montana	1983-1984
Nebraska	1983-1996
Nevada	1986-1996
New Hampshire	1984-1996
New Jersey	1984-1996
New York	1984-1996
North Carolina	1983-1996
North Dakota	1983-1996
Ohio	1984-1996
Oklahoma	1985-1996
Oregon	1983-1996
Pennsylvania	1984-1988, 1990-196
Rhode Island	1983-1987
South Carolina	1984-1996
South Dakota	1983-1986, 1991-1996
Tennessee	1983-1996
Texas	1983-1988, 1990-1992
Utah	1985-1996
Virginia	1985-1996
Washington	1983, 1985-1996
West Virginia	1983-1996
Wisconsin	1983-1996
Wyoming	1983-1986

Table 1. Summary Statistics

	Mean	Standard Deviation (Overall)	Standard Deviation of State Averages	Standard Deviation of Year Averages
Index Crimes per 100,000 pop	4819.31	1570.39	1429.42	454.83
Prison inmates per 1000 pop	1.97	1.64	1.15	0.96
Prison admissions per 1000 pop	1.44	0.91	0.87	0.36
Non-drug admissions per 1000 pop	1.11	0.62	0.60	0.21
Drug admissions per 1000 pop	0.32	0.35	0.33	0.15
Probationers per 100,000	748.69	483.81	380.24	234.28
Parolees per 100,000 pop	121.08	137.80	105.99	45.83
Jail inmates per 100,000 pop	111.28	77.70	57.95	42.62
% Population 15-17 (x100)	4.99	0.84	0.31	0.77
% Population 18-24 (x100)	11.78	1.62	0.63	1.44
% Population 25-34 (x100)	15.79	1.93	1.16	1.44
Unemployment Rate	6.41	2.07	1.27	1.22
% Black Population	10.52	12.21	12.30	0.35
% Metro Population	64.64	22.82	22.83	2.01
Mean Wage	-0.01	0.11	0.09	0.01
Mean Wage^2 (x100000)	1.13	2.34	1.79	0.22

There are 51 States (including Washington, DC) and 27 Years (1971-1997) for a total of 1377 observations. Due to missing data for some states and/or some years, there are only 1056 cases for Probationers per 100,000, only 1149 cases for Parolees per 100,000, only 1215 cases for Jail Inmates per 100,000, only 470 cases for Prison admissions per 1000, only 466 cases for Non-drug and Drug admissions per 1000, only 1263 cases for unemployment rate, and only 1317 cases for Mean Wage data.

Table 2. ML Regressions of Ln(Crime Rates/100,000):
Short or Longer Time Frame and with Basic or Extended Set of Control Variables.

	Model 1	Model 2	Model 3	Model 4
Intercept	7.747*** (0.193)	7.143*** (0.136)	7.245*** (0.209)	7.035*** (0.139)
Lag(SPP/1000 pop)	-0.007 (0.007)	0.025*** (0.004)	-0.018* (0.007)	0.007 (0.005)
% Population 15-17	11.971*** (1.991)	11.597*** (1.716)	7.079*** (2.004)	4.945*** (1.739)
% Population 18-24	-0.765 (0.627)	2.743*** (0.606)	2.473*** (0.686)	4.818*** (0.626)
% Population 25-34	1.245 (0.760)	2.121*** (0.520)	1.804* (0.759)	2.694*** (0.524)
Unemployment Rate			0.020*** (0.002)	0.016*** (0.002)
% Black Population			-0.037*** (0.007)	-0.040*** (0.005)
% Metro Population			0.012*** (0.003)	0.008*** (0.001)
Mean Wage			0.044 (0.053)	-0.108 (0.056)
Mean Wage^2 (x 100)			-1.441 (1.944)	-5.610* (2.216)
N (number of states)	51	51	51	51
T (number of years)	18	26	18	26
N Observations	917	1325	777	1185
-2 log likelihood	-2069.55	-2319.44	-1997.32	-2300.03
df (parameters estimated)	73	81	78	86

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

For all models, the first year of data is lost by taking the lag of Prison Population.
Model 1: Years 1971-1989, Basic Set of Control Variables
Model 2: Years 1971-1997, Basic Set of Control Variables
Model 3: Years 1971-1989, Extended Set of Control Variables
Model 4: Years 1971-1997, Extended Set of Control Variables

Table 3. ML Regressions of Ln(Crime Rates/100,000): Various Specifications.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	7.035*** (0.139)	7.296*** (0.133)	6.686*** (0.159)	7.088*** (0.316)	7.162*** (0.316)	7.335*** (0.316)	7.299*** (0.093)	7.161*** (0.099)
Lag(SPP/1000)	0.007 (0.005)		0.024*** (0.004)	0.004 (0.008)		0.011 (0.008)		
Lag[ln(SPP/1000)]		-0.166*** (0.018)			-0.037 (0.021)			
Ln(SPP/1000)							-0.082** (0.027)	-0.058* (0.027)
% Population 15-17	4.945*** (1.739)	3.855* (1.647)	4.336* (1.684)	-0.934 (3.297)	-1.401 (3.293)	-4.320 (2.408)	0.659 (1.157)	-0.133 (1.092)
% Population 18-24	4.818*** (0.626)	4.713*** (0.606)	4.409*** (0.632)	4.200*** (1.111)	4.160*** (1.110)	3.133*** (0.933)	3.080*** (0.428)	2.469*** (0.420)
% Population 25-34	2.694*** (0.524)	3.050*** (0.503)	2.215*** (0.521)	3.415*** (1.197)	3.369** (1.201)	2.871** (0.881)	0.426 (0.360)	0.109 (0.345)
Unemployment Rate	0.016*** (0.002)	0.017*** (0.002)	0.016*** (0.003)	0.006* (0.002)	0.006* (0.002)	0.005* (0.002)	0.004* (0.002)	0.002 (0.002)
% Black Population	-0.040*** (0.005)	-0.043*** (0.004)	-0.010*** (0.003)	0.001 (0.003)	0.003 (0.003)	0.001 (0.003)	-0.006 (0.003)	-0.002 (0.001)
% Metro Population	0.008*** (0.001)	0.006*** (0.001)	0.010*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.004*** (0.001)	0.004*** (0.001)

Mean Wage	-0.108 (0.056)	-0.108* (0.053)	-0.103 (0.057)	-0.004 (0.036)	-0.003 (0.036)	-0.001 (0.036)	0.005 (0.037)	0.015 (0.038)
Mean Wage^2 (x 100)	-5.610* (2.216)	-5.311* (-2.142)	-6.377*** (2.288)	1.352 (1.341)	1.380 (1.337)	1.364 (1.356)	-3.003* (1.510)	-2.602 (1.527)
Fixed Effects	State, Year	State, Year	None	Year	Year	None	State, Year	None
Random Effects	None	None	State, Year	State	State	State, Year	None	State, Year
AR(1)				Yes	Yes	Yes		
-2 log likelihood	-2300.03	-2379.85	-1820.22	-3149.46	-3151.65	-3020.69	-3225.89	-2819.29
df (parameters estimated)	86	86	13	37	37	13	88	15
BIC	-1691.4	-1771.2	-1728.2	-2887.6	-2889.8	-2928.7	-2603.1	-2713.1
AIC	-2128.0	-2207.9	-1794.2	-3075.5	-3077.7	-2994.7	-3049.9	-2789.3

For all models: N (number of states)=51; T (number of years)=26, Number of Cases=1185.
*** $p<0.001$ ** $p<0.01$ * $p<0.05$

Model 1: Years 1972-1997, Extended Set of Controls (same as Model 4 from Table 1).
Model 2: Years 1972-1997, Extended Set of Controls, Fixed Effects, lagged value of Ln(SPP) - elasticity specification.
Model 3: Years 1972-1997, Extended Set of Controls, Random Effects model
Model 4: Years 1972-1997, Extended Set of Controls, Fixed Year Effect, AR(1)
Model 5: Years 1972-1997, Extended Set of Controls, Fixed Year Effect, AR(1), lagged value of Ln(SPP) - elasticity specification
Model 6: Years 1972-1997, Extended Set of Controls, Random Effects, AR(1)
Model 7: Years 1972-1997, Extended Set of Controls, Fixed Effects, Error Correction Form
Model 8: Years 1972-1997, Extended Set of Controls, Random Effects, Error Correction Form

Table 4. ML Regressions of Ln(Crime Rates/100,000) with Additional Sanctions.

	Model 1	Model 2	Model 3	Model 4
Intercept	6.544*** (0.171)	6.417*** (0.177)	6.517*** (0.386)	7.307*** (0.264)
Lag(SPP/1000)	0.000 (0.006)	0.016** (0.005)	0.002 (0.009)	0.008 (0.009)
Lag(Probation/100,000)	0.000** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)
Lag(Parole/100,000)	0.000 (0.000)	0.000*** (0.000)	0.000 (0.000)	0.000 (0.000)
Lag(Jail/100,000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
% Population 15-17	6.180** (1.958)	5.839** (1.893)	2.955 (3.904)	-0.775 (2.809)
% Population 18-24	4.625*** (0.723)	3.168*** (0.688)	4.880*** (1.327)	2.435* (1.076)
% Population 25-34	5.497*** (0.669)	4.044*** (0.622)	5.549*** (1.580)	2.175* (1.077)
Unemployment Rate	0.011*** (0.003)	0.006* (0.003)	0.009** (0.003)	0.007* (0.003)
% Black Population	-0.031*** (0.007)	-0.006* (0.003)	0.000 (0.003)	0.000 (0.003)

% Metro Population	0.010*** (0.001)	0.011*** (0.001)	0.007*** (0.001)	0.007*** (0.001)
Mean Wage	-0.218** (0.075)	-0.165* (0.077)	-0.050 (0.049)	-0.033 (0.050)
Mean Wage^2 (x100)	-18.269*** (4.281)	-20.091*** (4.438)	4.207 (2.925)	4.479 (2.971)
Fixed Effects	Yes	No	Year Only	No
-2 log likelihood	-1840.58	-1461.94	-2289.17	-2197.85
df (parameters estimated)	77	16	34	16
BIC	-1318.5	-1353.5	-2058.7	-2089.4
AIC	-1686.6	-1429.9	-2221.2	-2165.8

For all models: N (number of states)=45; T (number of years)=20, Number of Cases=880, Inclusion of AR(1) error specification.

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

Model 1: Fixed Effects with Additional Sanctions
Model 2: Random Effects with Additional Sanctions
Model 3: Fixed Year Effect, AR(1), and Additional Sanctions
Model 4: Random Effects, AR(1), and Additional Sanctions

Table 5. ML Regressions of Ln(Crime Rates/100,000) with Additional Sanctions and Specifications.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	6.562^{***} (0.387)	7.406^{***} (0.279)	4.973^{***} (0.521)	5.638^{***} (0.466)	6.523^{***} (0.394)	7.312^{***} (0.262)
Lag(SPP/1000)	-0.026 (0.017)	-0.007 (0.016)	-0.011 (0.011)	-0.009 (0.011)	various#	0.010# (0.009)
Lag($(SPP/1000)^2$)	0.002 (0.001)	0.001 (0.001)	—	—	—	—
Lag(Prob/100,000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Lag(Parole/100,000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Lag(Jail/100,000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
% Population 15-17	2.657 (3.903)	-1.479 (2.879)	2.195 (3.983)	5.576 (3.470)	2.474 (1.632)	-0.834 (2.735)
% Population 18-24	5.231^{***} (1.338)	2.400^{**} (1.076)	4.809^{***} (1.293)	4.461^{***} (1.272)	5.424^{***} (1.397)	2.492^{*} (1.063)
% Population 25-34	5.526^{***} (1.582)	1.911 (1.104)	8.756^{***} (1.528)	6.620^{***} (1.357)	5.125^{**} (1.591)	2.118^{*} (1.059)
Unemployment Rate	0.009^{**} (0.003)	0.006^{*} (0.003)	0.008^{**} (0.003)	0.007^{**} (0.003)	0.009^{**} (0.003)	0.006^{*} (0.003)

% Black Population	0.001 (0.003)	0.001 (0.003)	-0.006* (0.003)	-0.006* (0.003)	0.001 (0.003)	0.000 (0.003)
% Metro Population	0.007*** (0.001)	0.007*** (0.001)	0.010*** (0.001)	0.010*** (0.001)	0.007*** (0.001)	0.007*** (0.001)
Mean Wage	-0.049 (0.049)	-0.032 (0.050)	-0.051 (0.050)	-0.038 (0.050)	-0.039 (0.049)	-0.032 (0.050)
Mean Wage^2	4.188 (2.914)	4.476 (2.966)	3.514 (2.951)	3.915 (2.984)	4.542 (2.885)	4.468 (2.962)
Fixed Effects	Yes	No	Yes	No	Yes	No
Random Effects	No	Yes	No	Yes	No	Yes
Quadratic SPP	Yes	Yes	No	No	No	No
State-specific Time trend	No	No	Yes	Yes	No	No
Varying SPP coefficient over time	No	No	No	No	Yes (Fixed)	Yes (Random)
-2 log likelihood	-2292.83	-2198.97	-2356.13	-2265.17	-2318.91	-2200.53
df (parameters estimated)	35	17	77	61	53	18
BIC	-2055.5	-2083.7	-1834.1	-1851.6	-1959.6	-2078.5
AIC	-2222.8	-2165.0	-2202.1	-2143.2	-2212.9	-2164.5

For all models: N (number of states)=45; T (number of years)=20, Number of Cases=880, Inclusion of AR(1) error specification.

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

Model 1: Fixed Year Effect, Additional Sanctions, Quadratic SPP
Model 2: Random Year Effect, Additional Sanctions, Quadratic SPP
Model 3: Fixed Year Effect, Additional Sanctions, State-specific time trend
Model 4: Random Year Effect, Additional Sanctions, State-specific time trend
Model 5: Fixed Year Effect, Additional Sanctions, Fixed Time-varying SPP coefficient
Model 6: Random Year Effect, Additional Sanctions, Random Time-varying SPP coefficient

[#] For magnitudes of the time-varying coefficient, see Table 6.

Table 6. Magnitude of the time-varying coefficient of prison population for Model 5and Model 6 in Table 5

Year	Effect from Table 5, Model 5	Effect from Table 5, Model 6	Total effect from Table 5, Model 6#
1978	-0.0593	0.0044	0.0145
1979	-0.0589	-0.0018	0.0083
1980	-0.0455	-0.0061	0.0040
1981	-0.0423	-0.0050	0.0051
1982	-0.0247	0.0009	0.0110
1983	-0.0496	0.0049	0.0150
1984	-0.0441	0.0072	0.0173
1985	-0.0463	0.0037	0.0138
1986	-0.0382	0.0009	0.0110
1987	-0.0388	-0.0009	0.0092
1988	-0.0165	0.0003	0.0104
1989	-0.0125	-0.0010	0.0091
1990	-0.0123	-0.0025	0.0076
1991	-0.0099	-0.0025	0.0076
1992	-0.0072	-0.0014	0.0087
1993	0.0022	0.0015	0.0116
1994	0.0003	-0.0009	0.0092
1995	0.0029	-0.0015	0.0086
1996	0.0084	0.0020	0.0121
1997	0.0053	-0.0023	0.0124

#Total effect is formed by adding the fixed overall mean of 0.010 reported in Model 6, Table 5 for lag(SPP/1000) to the random effects reported in column two above.

Table 7. ML Regressions of Ln(Crime Rates/100,000): SPP Admissions Flow, Various Specifications.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	6.576*** (0.214)	6.597*** (0.213)	6.501*** (0.228)	6.058*** (0.333)	6.062*** (0.333)	7.578*** (0.232)	6.994*** (0.146)	7.225*** (0.105)
Lag(Admissions/1000)	-0.007 (0.012)		0.022 (0.012)	0.004 (0.009)		0.014 (0.009)		
Ln(Admissions/1000)							-0.003 (0.019)	0.018 (0.020)
Lag[ln(Admissions/1000)]		-0.037 (0.022)			-0.008 (0.016)		-0.035 (0.018)	-0.009 (0.019)
% Population 15-17	0.090 (2.400)	0.189 (2.372)	0.559 (2.533)	6.401 (3.304)	6.214 (3.302)	-2.953 (2.886)	3.050 (1.586)	4.153** (1.490)
% Population 18-24	5.726*** (0.972)	5.855*** (0.969)	4.747*** (0.972)	5.532*** (1.181)	5.697*** (1.179)	0.993 (0.981)	2.900*** (0.725)	0.287 (0.592)
% Population 25-34	7.058*** (0.882)	6.977*** (0.876)	5.286*** (0.868)	7.723*** (1.357)	7.677*** (1.359)	2.347* (0.942)	1.824** (0.629)	1.023* (0.477)
Unemployment Rate	0.012*** (0.004)	0.013*** (0.004)	0.008* (0.004)	0.011*** (0.003)	0.011*** (0.003)	0.006* (0.003)	-0.003 (0.002)	-0.009*** (0.002)
% Black Population	-0.042*** (0.009)	-0.041*** (0.009)	-0.001 (0.003)	0.002 (0.002)	0.002 (0.003)	0.004 (0.003)	0.011 (0.007)	0.001 (0.001)
% Metro Population	0.009*** (0.002)	0.008*** (0.002)	0.008*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.006*** (0.001)	0.004* (0.001)	0.002*** (0.001)

Mean Wage	-0.055 (0.079)	-0.054 (0.079)	-0.021 (0.083)	-0.104* (0.043)	-0.106* (0.043)	-0.073 (0.045)	0.097 (0.053)	0.142** (0.054)
Mean Wage^2 (x 100)	-14.920** (5.187)	-15.692** (-5.174)	-13.963* (5.513)	-3.025 (2.897)	-2.960 (2.897)	-2.735 (3.039)	-0.534 (3.588)	1.404 (3.802)
Fixed Effects	State, Year	State, Year	None	Year	Year	None	State, Year	None
Random Effects	None	None	State, Year	State	State	State, Year	None	State, Year
AR(1)				Yes	Yes	Yes		
-2 log likelihood	-1288.30	-1290.81	-951.60	-1558.93	-1559.02	-1467.90	-1565.95	-1320.06
df (parameters estimated)	65	65	13	25	25	13	65	15
BIC	-888.4	-890.9	-871.6	-1405.1	-1405.2	-1387.9	-1173.6	-1229.5
AIC	-1158.3	-1160.8	-925.6	-1508.9	-1509.0	-1441.9	-1436.0	-1290.1

For models 1 to 6: N (number of states)=42; T (number of years)=14, Number of Cases=470.
For models 7 and 8: N (number of states)=42; T (number of years)=13, Number of Cases=418.

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

Model 1: Years 1984-1997, Extended Set of Controls.
Model 2: Years 1984-1997, Extended Set of Controls, Fixed Effects, lagged value of Ln(SPP) - elasticity specification.
Model 3: Years 1984-1997, Extended Set of Controls, Random Effects model
Model 4: Years 1984-1997, Extended Set of Controls, Fixed Year Effect, AR(1)
Model 5: Years 1984-1997, Extended Set of Controls, Fixed Year Effect, AR(1), lagged value of Ln(SPP) - elasticity specification
Model 6: Years 1984-1997, Extended Set of Controls, Random Effects, AR(1)
Model 7: Years 1984-1996, Extended Set of Controls, Fixed Effects, Error Correction Form
Model 8: Years 1984-1996, Extended Set of Controls, Random Effects, Error Correction Form

Table 8. ML Regressions of Ln(Crime Rates/100,000): SPP NonDrug and Drug Admissions Flow, Various Specifications.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	6.597*** (0.211)	6.589*** (0.210)	6.467*** (0.227)	6.136*** (0.330)	6.128*** (0.332)	7.636*** (0.214)	7.045*** (0.146)	7.257*** (0.105)
Lag(NonDrug Admissions/1000)	0.047** (0.018)		0.076*** (0.018)	-0.016 (0.011)		-0.013 (0.012)		
Ln(NonDrug Admissions/1000)							0.018 (0.019)	0.017 (0.020)
Lag[ln(NonDrug Admissions/1000)]		0.056** (0.024)			-0.021 (0.016)		-0.022 (0.018)	-0.010 (0.020)
Lag(Drug Admissions/1000)	-0.109*** (0.030)		-0.094** (0.031)	0.070** (0.026)		0.110*** (0.026)		
Ln(Drug admissions/1000)							-0.011 (0.011)	0.002** (0.012)
Lag[ln(Drug admissions/1000)]		-0.050*** (0.011)			0.013 (0.010)		-0.003 (0.011)	0.001 (0.011)
Lag[ln(Index Crimes/100,000)]							0.000*** (0.000)	0.000*** (0.000)
% Population 15-17	-1.746 (2.386)	-1.264 (2.329)	-1.275 (2.508)	6.816* (3.284)	6.268 (3.305)	-2.606 (2.681)	2.308 (1.574)	3.816* (1.501)
% Population 18-24	6.175*** (0.967)	6.070*** (0.949)	5.175*** (0.960)	4.519*** (1.200)	5.163*** (1.184)	-0.460 (0.926)	2.862*** (0.718)	0.381 (0.586)

% Population 25-34	6.941*** (0.874)	6.993*** (0.866)	5.393*** (0.863)	7.791*** (1.342)	7.683*** (1.351)	3.069*** (0.868)	1.521* (0.640)	0.916 (0.478)
Unemployment Rate	0.012*** (0.003)	0.009** (0.003)	0.008* (0.004)	0.010*** (0.003)	0.011*** (0.003)	0.005 (0.003)	-0.004 (0.002)	-0.009*** (0.002)
% Black Population	-0.034*** (0.009)	-0.031*** (0.009)	-0.000 (0.003)	0.002 (0.002)	0.002 (0.003)	0.003 (0.003)	0.015* (0.007)	0.001 (0.001)
% Metro Population	0.010*** (0.002)	0.008*** (0.002)	0.008*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.006*** (0.001)	0.004** (0.001)	0.002*** (0.001)
Mean Wage	-0.029 (0.078)	-0.046 (0.077)	0.009 (0.082)	-0.109* (0.043)	-0.106* (0.043)	-0.084 (0.044)	0.105* (0.052)	0.145** (0.053)
Mean Wage^2 (x 100)	-14.156** (5.092)	-15.513** (-5.065)	-13.607* (5.375)	-2.711 (2.883)	-3.056 (2.896)	-2.438 (3.031)	-1.318 (3.540)	0.833 (3.746)
Fixed Effects	State, Year	State, Year	None	Year	Year	None	State, Year	None
Random Effects	None	None	State, Year	State	State	State, Year	None	State, Year
AR(1)				Yes	Yes	Yes		
-2 log likelihood	-1298.98	-1304.56	-965.36	-1550.24	-1545.11	-1466.32	-1570.64	-1322.59
df (parameters estimated)	66	66	14	26	26	14	67	17
BIC	-893.5	-899.0	-879.3	-1390.5	-1385.4	-1380.3	-1166.9	-1220.1
AIC	-1167.0	-1172.6	-937.4	-1498.2	-1493.1	-1438.3	-1436.6	-1288.6

For models 1 to 6: N (number of states)=42; T (number of years)=14, Number of Cases=466.

For models 7 and 8: N (number of states)=41; T (number of years)=13, Number of Cases=414.

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

Model 1: Years 1984-1997, Extended Set of Controls.
Model 2: Years 1984-1997, Extended Set of Controls, Fixed Effects, lagged value of Ln(SPP) - elasticity specification.
Model 3: Years 1984-1997, Extended Set of Controls, Random Effects model
Model 4: Years 1984-1997, Extended Set of Controls, Fixed Year Effect, AR(1)
Model 5: Years 1984-1997, Extended Set of Controls, Fixed Year Effect, AR(1), lagged value of Ln(SPP) - elasticity specification
Model 6: Years 1984-1997, Extended Set of Controls, Random Effects, AR(1)
Model 7: Years 1984-1996, Extended Set of Controls, Fixed Effects, Error Correction Form
Model 8: Years 1984-1996, Extended Set of Controls, Random Effects, Error Correction Form

Table 9. ML Regressions of Ln(Crime Rates/100,000): SPP NonDrug and Drug Admissions Flow and Sentencing Practices, Various Specifications.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	5.023*** (0.412)	4.887*** (0.398)	7.312*** (0.222)	6.032*** (0.459)	5.967*** (0.461)	7.535*** (0.230)	6.035*** (0.295)	7.372*** (0.145)
Lag(NonDrug Admissions/1000)	-0.030 (0.021)		0.024 (0.023)	-0.004 (0.014)		0.009 (0.016)		
Ln(NonDrug Admissions/1000)							-0.006 (0.024)	-0.009 (0.027)
Lag[ln(NonDrug Admissions/1000)]		-0.080** (0.029)			-0.020 (0.023)		-0.066** (0.023)	-0.039 (0.026)
Lag(Drug Admissions/1000)	-0.038 (0.030)		-0.037 (0.042)	-0.047 (0.043)		0.010 (0.045)		
Ln(Drug admissions/1000)							-0.046*** (0.014)	-0.020 (0.015)
Lag[ln(Drug admissions/1000)]		-0.027 (0.011)			0.002 (0.014)		0.022 (0.013)	0.026 (0.014)
Lag[ln(Index Crimes/100,000)]							0.000*** (0.000)	0.000*** (0.000)
Sentencing Practices x 100	0.009 (0.018)	0.001 (0.017)	0.027 (0.021)	0.017 (0.014)	0.017 (0.014)	0.022 (0.015)	0.015 (0.012)	0.014 (0.014)
% Population 15-17	5.914 (3.196)	6.661* (2.329)	-3.943 (2.830)	-3.923 (3.964)	-2.719 (3.941)	-6.768* (2.712)	3.843 (2.202)	-0.009 (2.002)

% Population 18-24	7.949*** (1.037)	7.959*** (0.997)	3.707*** (1.014)	8.266*** (1.529)	7.858*** (1.549)	1.691 (1.124)	5.740*** (0.745)	2.404*** (0.703)
% Population 25-34	9.845*** (1.658)	11.062*** (1.634)	1.117 (0.983)	7.895*** (2.264)	8.311*** (2.250)	2.006 (1.043)	3.918** (1.259)	-0.632 (0.706)
Unemployment Rate	0.005 (0.003)	0.005 (0.004)	-0.005 (0.004)	0.011** (0.004)	0.011** (0.004)	0.002 (0.004)	0.002 (0.003)	-0.007* (0.003)
% Black Population	-0.038* (0.016)	-0.042* (0.016)	0.006* (0.003)	0.003 (0.003)	0.002 (0.003)	0.006 (0.003)	0.017 (0.012)	0.000 (0.001)
% Metro Population	0.015*** (0.003)	0.013*** (0.003)	0.010*** (0.001)	0.008*** (0.002)	0.008*** (0.002)	0.009*** (0.002)	0.004 (0.002)	0.003** (0.001)
Mean Wage	-0.071 (0.093)	-0.015 (0.089)	0.037 (0.107)	-0.123 (0.066)	-0.123 (0.065)	-0.078 (0.071)	-0.002 (0.062)	0.040 (0.071)
Mean Wage^2 (x 100)	0.610 (6.880)	2.815 (6.607)	-3.394 (8.209)	2.764 (4.731)	2.810 (4.744)	2.046 (5.143)	2.733 (4.709)	3.885 (5.477)
Fixed Effects	State, Year	State, Year	None	Year	Year	None	State, Year	None
Random Effects	None	None	State, Year	State	State	State, Year	None	State, Year
AR(1)				Yes	Yes	Yes		
-2 log likelihood	-829.82	-847.54	-562.47	-781.63	-780.80	-728.88	-1017.00	-781.81
df (parameters estimated)	57	57	15	26	26	15	60	18
BIC	-516.3	-534.0	-480.0	-638.6	-637.8	-646.4	-687.2	-682.9
AIC	-715.8	-733.5	-532.5	-729.6	-728.8	-698.9	-897.0	-745.8