Discussion of William Sabol’s paper “Implications of Criminal Justice System Adaptation for Prison Population Growth and Corrections Policy.”

By Steve Aos, Director, Washington State Institute for Public Policy


Dr. Sabol’s paper offers readers a “twofer”—it covers two areas in one paper.

The first part of the paper concentrates on decomposing trends in state prison incarceration levels between 1994 and 2006. Following and extending the work of Blumstein and Beck (2005), the paper presents useful information on criminal justice system factors that appear to be associated with changes in incarceration levels.

We learn, for example, that system factors have changed over time, as well as by type of crime. In more recent years, increases in the length of prison stay have been more important determinants of prison population growth for violent offenses than for property or drug offenses. We also learn that the ability of prosecutors to turn arrests into convictions has increased over time for violent and property offenses and this is correlated with increased incarceration rates. And arrests made by police for drug offenses have been associated with changed incarceration levels.

The implications of these decomposed trends—if they are causally related to incarceration levels—present a daunting challenge to policy making. Not only are there 50 state-level policy experiments going on in the United States, but Sabol’s results suggest that there are a multitude of sub-state aspects of the criminal justice system involving prosecutors, defenders, courts, and police. Just as it can be hard to direct a play with many actors, policy wise it can be difficult to direct system change from state capitals (let alone from the nation’s capital) when there are so many ostensibly independent criminal justice institutions on the policy stage. If you push policy levers at one point in the system, another part of the system may pull in the contrary direction. This “push-pull” factor may or may not be causal, but Sabol’s analysis indicates that it at least appears to evolve over time.

The research implications from the first part of paper indicate that we need to know a lot more about the internal workings of the amoeba-like system if the policy goal is to bring about more cost-beneficial statewide criminal justice policies. We need to know, for example, how police, prosecutors, and courts respond to changes in state legislative and executive sentencing-related policies, and vice versa. And, since heterogeneity among all of these actors throughout the United States is likely to be the rule not the exception, we need to know about the variation in these

relationships at the sub-national detail. That is, county prosecutors in Washington State are likely to respond to policy changes from Olympia in different ways than prosecutors in Los Angeles respond to policy changes in Sacramento.

In the second part of his paper, Sabol begins by noting that the first part of the paper “is severely limited.” He notes that this limitation surfaces because the decomposition analysis presented in the first part “cannot address causal issues related to prison population growth even if it can point out areas for consideration and research.” Bummer.

The paper proceeds by providing a brief and selected narrative review of some studies that have, in turn, summarized the research literature on the causal relationships between crime and criminal justice policies. From a policy perspective, the news here is not pretty. These reviews indicate that, in general, researchers have been unable to contribute precise and convincing estimates of particular policy options for several reasons. First, researchers in this area have been forced to use fairly aggregated data because they usually lack individual-level information at the point where actual decisions are made by criminals and criminal justice practitioners. With so much information unobserved (to researchers), and with random assignment the exception not the rule, it has been difficult to model relationships credibly.

The most significant reason for a weak research base, however, is that criminal justice policies give every known indication of being endogenous. Crime and criminal justice policies appear to be determined jointly and this has confounded the ability of researchers to identify causal relationships that would otherwise be so useful for setting policies. Sabol notes that the usual statistical procedures to overcome endogeneity—unassailable instrumental variables—are very hard to come by in a world where so many things are interrelated. In concluding the second part of the paper, Sabol encourages the research community to continue to look for better instruments, and that it should also look for try to exploit natural experiments and employ matching estimators.

On this last point, it is worth emphasizing that, if more widely used, regression discontinuity approaches may be able to help identify some causal effects for certain sentencing-related policies. When states, such as Washington State, implement presumptive statewide sentencing grids, they are forced to adopt somewhat arbitrary cut-points on the particular sanctions received by convicted offenders. Regression discontinuity designs offer the potential to at least identify “local treatment” effects around some of these sentencing rules. And, since sentencing grids have many of these cutoff levels embedded in them, the estimation of many local treatment effects offer the potential to inform a more generalized set of estimates about sentencing. This is a worthy area of future research, particularly because states with sentencing grids may often have comprehensive administrative databases with sentencing and conviction data (because the sanctions for an offender’s subsequent offenses will typically be a function the current offense and criminal history). Thus an important part of a national research agenda could focus on regression discontinuity designs in those states with grid systems and good electronic administrative data.
Research Needs: More Fully Developing the Current Model

Springing off the concepts addressed in Sabol’s paper, the remainder of my comments centers on what I think are the most pressing needs for research. I offer these suggestions from my position as a policy advisor to the Washington State legislature. While there is a host of interesting research questions raised by the paper, I only focus on those that I think are most clearly tied to the practical decisions that face legislative bodies.

Legislatures need, I believe, two things from researchers to help them form more effective criminal justice public policies: consistently estimated “crime impact statements” and “fiscal impact statements” for a portfolio of policy choices. This information would allow the computation of costs and benefits, return on investment, and the riskiness of portfolios of policy choices.

I start with the assumption that legislatures want to accomplish three principal goals for each criminal justice policy they consider: improve justice, reduce crime (victim costs), and control taxpayer costs. The agenda I discuss here is only related to the last two of the troika: crime and taxpayer spending. I do not address a research agenda for the justice sentencing goal.

To what degree, then, can the current state of research help states create a crime impact statement and a fiscal impact statement so that current policies can be analyzed? As it turns out, there is a fairly large research literature on the effect of incarceration rates on crime. Unfortunately, this research lacks the detail necessary to inform very many specific decisions that typically arise in state capitals. If we highlight the limitations in this research, it can point to areas in need of improvement.

Many of the existing studies addressing the prison-crime relationship in the United States construct models using state-level data over a number of years to estimate the parameters of an equation of this general form:

\[
\text{In this typical model, crime } C \text{ in state } s \text{ and year } y \text{ is estimated to be a function of a state’s overall average daily prison population, } ADP, \text{ a vector of control variables, } X, \text{ and an error term, } e. \text{ Crime is most frequently measured with data from the FBI’s Uniform Crime Reporting (UCR) system. The variables are usually divided by population so that they are expressed as rates. The models are almost always estimated with a log-log functional form, at least for the dependent and the main “independent” policy variables. Several authors have also observed that the state-level time series data often used to estimate equation (1) are likely have unit roots. Thus, to help avoid estimating spurious relationships, some authors estimate equation (1) in first differences since the time series data...}
\]

---


\[3\text{See, for example, Marvell, 2009. See also, W. Spelman (2008). Specifying the relationship between crime and prisons, Journal of Quantitative Criminology, 24, 149-178.}\]
typically do not exhibit unit roots after differencing once. As noted above, there is also considerable concern in the research literature on the econometric implication of the possible simultaneous relationship between the variables of interest in equation (1): that is, crime may be a function of ADP, but ADP may also be a function of crime. This simultaneous relationship can cause statistically biased estimates if not dealt with.

In order to produce a “crime impact statement,” marginal effects from this generic log-log crime model are then obtained with:

\[
\text{In equation (2), the change in crime is estimated with } E, \text{ the crime-prison elasticity obtained from coefficient } b \text{ of the typical loglog estimation of equation (1); UCR, the reported crime rate; ADP, the incarceration rate, and RRate, the reporting rate to police by crime victims. The marginal effects are sometimes calculated either at the mean values for } ADP{\text{-UCR}}{\text{-RRates}} \text{ or, more to the point for policy purposes, at the most recent values for } ADP{\text{UCR}}{\text{-RRates}}. \text{The log-log estimation of the constant elasticity } E \text{ implies diminishing returns when } E \text{ is less than one and incarceration rates are raised. Similarly, an elasticity less than one coupled with reduced ADP implies increasing returns.}
\]

To understand the needs for improved research, it is instructive to examine current limitations in both the data and methodological assumptions in equation (1).

We can begin with the dependent variable, crime. In the studies estimating these types of equations, crime is most often measured with data from the FBI’s UCR. These data count the number of crimes reported to police. Some studies estimate a model of total UCR crime reported to police, while other studies estimate two equations, one for violent crime reported to police, and another for property crime reported to police. Still other studies break the analysis down further and estimate equations for the seven major types of “Part 1” crimes in the UCR data: murder, rape, robbery, aggravated assault, burglary, theft, and motor vehicle theft.

One significant problem with the “Part 1” UCR crime data is that they do not match directly how some states, including Washington, define felony crimes. In Washington, this applies to two types of crimes in particular: felony sex crimes and theft/larceny. The UCR sex offense data only count rapes of females over the age of 12. In addition to this obvious limitation in the UCR data, there are other felony sex crimes (e.g., child molestation), defined by the Revised Code of Washington that are not included in the UCR rape category. Similarly, the UCR data count some types of theft crimes that are below the threshold of felony theft in Washington. Therefore, adjustments to the UCR data are needed to better match the data needs for actual policy making in the states. A comprehensive state-by-state review of criminal codes and ways to adjust the UCR data needs to be undertaken.

We can then examine the main policy variable, average daily prison population. In virtually all studies in the current research literature, the policy variable examined is statewide prison average.
daily population. In most studies, this is measured by counting the total number of inmates at the beginning of a year, or the end of a year, and dividing by a state’s population aggregate to obtain an overall incarceration rate.

Measuring ADP with the total number of offenders—as opposed to more refined categories of offenders convicted of violent, property, or drug offenses, or defining offenders based on an actuarial risk assessment as high-risk, moderate-risk, or lower-risk offenders—has been necessary in cross-state analyses, because total ADP is usually the only information available. This restriction, however, means that the typical research study only measures the average effect on crime of the average offender sentenced to prison. The criminal propensities of different types of offenders, on the other hand, are quite heterogeneous in terms of the amount and types of crime committed. It is widely known, for example, that sex offenders have different crime characteristics than property offenders. We know that there is wide heterogeneity in offenders with respect to the probability of re-offense. Yet, equation (1) treats all ADP as if it is a homogeneous unit of measurement. This severely limits the ability of current research findings to offer practical policy advice when crafting specific sentencing policies.

Given trends in sentencing policies in the United States, this “average offender” limitation poses at least three empirical problems. First, the average mix of offenders in prison has changed over time. For example, in Washington State, there were virtually no offenders in prison for drug crimes prior to the mid-1980s. Sentencing laws were changed in the late 1980s and the average proportion of drug offenders in ADP increased substantially. The average risk for reoffense has also exhibited long term trends. Among offenders released from prison in Washington, there has been a 23 percent increase in offenders’ risk level between 1991 and 2005. Thus, the average crime/ADP coefficient from most regressions may not be aligned with the current mix of offenders in a state’s ADP.

The second reason why parameters in models like equation (1) are limited in their ability to inform actual policy choices facing legislatures is that policy decisions to raise or lower ADP are not usually across-the-board or “average” decisions. A legislature will rarely raise sentences for all types of crimes by a uniform amount, nor will a legislature typically lower sentences uniformly for all types of crimes (although this has been done). If a legislature were to uniformly lower lengths of stay, for example, a high risk sex offender would be treated the same as a low risk drug offender. Since this is likely to be seen as an undesirable policy, if prison ADP is to be adjusted, legislative discussions are more likely to focus on at least some level of policy selectivity in which types of offenders are released early. Much more often, a legislature will adjust sentencing statutes for particular types of crimes, rather than adopt across-the board changes.

The third significant reason why the average elasticity estimates from current research in not too useful for policy making is that not all policies that affect prison ADP have an equal effect on crime.

---


5 Ibid.
Durlauf and Nagin (2010) provide a very useful review and analysis of the research literature on the two sentencing factors that determine a state’s ADP: the probability of a sentence to prison given a conviction, and the severity of the sentence in terms of length of prison stay.

Each of these sentencing parameters—the certainty of punishment and the severity of punishment—are affected by different sentencing policies. Yet, as Durlauf and Nagin found, the research literature indicates that the two types of policies are likely to have quite different effects on crime. They conclude:

*The key empirical conclusion of our literature review is that there is relatively little reliable evidence of variation in the severity of punishment having a substantial deterrent effect but that there is relatively strong evidence that variation in the certainty of punishment has a large deterrent effect.*

Using the Durlauf and Nagin results, one would conclude that the mean ADP elasticity from equation (1) for a sentencing policy that affects the certainty of punishment would be higher. Conversely, the mean ADP elasticity for a sentencing policy that affects the length of prison stay would be lower. While the state of research may not allow a clear delineation of the magnitude of these differential effects, the direction seems clear based on the findings of Durlauf and Nagin.

This means that the coefficients obtained from equations such as (1) above can be thought of as only rough guides for the effectiveness of average sentencing changes. Until new research can address the issues raised, the coefficients obtained from these equations need to be adjusted to better estimate the specific policy choices available to legislatures. Adjustments need to reflect: (a) the wide heterogeneity of criminal propensities among offenders, (b) that legislatures usually adjust sentencing policies differentially for different types of crimes, and (c), that the particular type of sentencing policy adopted is likely to affect crime differentially depending on whether the policy changes the certainty or severity of punishment.

Finally, it is worth re-emphasizing the main concluding point made by Sabol: the inability of most current research to control for the apparent simultaneity between crime and criminal justice policies. Crime may be affected by prison, but there is also evidence in many of the studies that the use of prison is affected by crime. This simultaneous relationship, if not accounted for, will probably bias the coefficient in a model like equation (1) downward. If a legislature’s willingness to provide prison cells is motivated by changes in crime levels, then the observed relationship between prison and crime can be measuring both prison supply decisions and criminal response to prison levels. Therefore, an observed effect of prison on crime is likely to be muted because some of the observed relationship reflects the use of more prison as a result of crime changes. As Sabol notes, technically these models require an exogenous source of variation—an instrumental variable

---


7 Marvell, 2009.
or a discontinuity around some arbitrary sentencing cut-off level—that affects the use of prison but is probably otherwise unrelated to the error term in equation (1). These instrumental variables are hard to find so there are many more estimates that do not account for simultaneity than that do. Future research needs to address this limitation. In the existing research literature, there have been only a few attempts to measure credibly the magnitude of this simultaneous relationship.\(^8\)