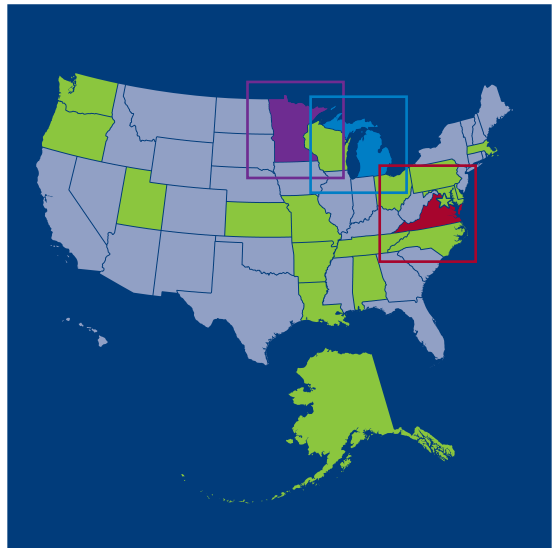


Assessing Consistency and Fairness in Sentencing: *A Comparative Study in Three States*



■ ■ ■ States with Sentencing Guidelines Systems



Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

Authors:

National Center for State Courts

Brian J. Ostrom
Charles W. Ostrom
Roger A. Hanson
Matthew Kleiman

Information design:

VisualResearch, Inc.

Neal B. Kauder

Design and layout:

Mazmedia

Michael A. Zanconato

Production of this report was funded by the Public Safety Performance Project of the Pew Charitable Trust's Center on the States.

Launched in 2006, the Public Safety Performance Project seeks to help states advance fiscally sound, data-driven policies and practices in sentencing and corrections that protect public safety, hold offenders accountable, and control corrections costs.

The Pew Charitable Trusts applies the power of knowledge to solve today's most challenging problems.

Pew's Center on the States identifies and advances effective policy approaches to critical issues facing states.

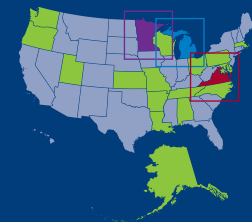
www.pewcenteronthestates.org

This report summarizes material from a larger study produced with support from the National Institute of Justice (2003-IJ-CX-1015). The authors gratefully acknowledge the generous support of NIJ and the encouragement of Linda Truitt, our project monitor.



The mission of the National Center for State Courts is to improve the administration of justice through leadership and service to state courts and courts around the world.





Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

Criminal sentencing in the American states has undergone substantial changes during the past several decades. A major policy shift affecting many offenders is the introduction of structured sentencing. Policies popularly known as three strikes, truth-in-sentencing, and mandatory minimum imprisonment have taken hold in some states, but a more widespread, substantial legal policy is the introduction of sentencing guidelines in at least 20 states and the District of Columbia.

Sentencing guidelines are a relatively new reform effort to encourage judges to take specific legally relevant elements into account in a fair and consistent way when deciding whether a convicted offender should be imprisoned, and if so, for what length of time. A common concern of state policymakers for limiting sentencing disparity under indeterminate sentencing laws is a fundamental rationale for the adoption of guidelines. For this reason, most states make explicit reference in their statement of purpose to achieving the goals of consistency (predictability and proportionality) and fairness (non-discrimination) in sentencing. Exploring how well alternative guideline systems realize these twin goals is the aim of the current research.

Guidelines consist of two main parts

- (1) A specified set of elements to be considered, such as the formal nature of the conviction offense and the offender's past criminal history.
- (2) Instructions on how the elements are to be weighted or scored in terms of their gravity.

Because both the elements and the corresponding mechanics of how they are to be applied in individual cases are highly differentiated and nuanced, guideline systems vary considerably from state to state. Comparisons among guidelines are often couched in the language of one system being more or less “mandatory” or “voluntary” than another. For example, stricter departure policies, tighter sentencing ranges, and more vigorous appellate review are aspects of what are usually called more mandatory, in the sense of being presumptive, systems.

In contrast, under a voluntary (or advisory) guideline system, judges are not required to follow a particular sentencing recommendation, but must usually provide a reason when the recommendation is not followed. Implicit in a preference for more mandatory versus more voluntary guidelines, is a judgment on the degree to which judicial discretion must be constrained to best achieve consistency and fairness.

Key policy questions

- (1) Have states designed sentencing guidelines that achieve a high level of predictability without denying judges adequate discretion in each individual case?
- (2) Are there important similarities or differences in sentencing patterns among states with different guideline structures and organization?
- (3) What lessons can be drawn from the experiences in Minnesota, Michigan and Virginia for other states around the country?



Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

Executive Summary

The National Center for State Courts conducted an in-depth examination of sentencing patterns in three states with substantially different guidelines systems:

Minnesota, which has a relatively strict system;

Michigan, whose guidelines offer more judicial discretion, and

Virginia, where compliance with the recommended sentences is completely voluntary.

Ultimately, how one interprets the observed differences in outcomes among the three states will reflect individual views on the appropriate level of judicial discretion.

At the conceptual level, desired consistency in sentencing outcomes clashes with desirable judicial discretion because they involve quite different fundamental assumptions. On the one hand, consistency posits that the most relevant criteria for classifying cases are identifiable and applicable to all cases. On the other hand, discretion posits that cases are sufficiently different to make it nearly impossible to establish a common means of comparison in each individual case. This study accepts the creative tension between consistency and discretion, which seems reasonable given the current state of knowledge, and therefore makes no attempt to rank the overall effectiveness of the three systems.

Because all guideline systems reflect alternative choices about the appropriate level of judicial discretion, the study identifies six criteria that define and distinguish sentencing guideline systems in the United States.

Based on the how each state to these criteria, the study places all of the existing guideline systems along a continuum from more voluntary to more mandatory. From this perspective, it is possible to assess the degree to which three important sentencing goals – predictability, proportionality, elimination of discrimination – are realized in the context of sentencing systems at various points on the sentencing continuum. This study selected the three states of Minnesota, Michigan and Virginia because they fall at different points along the continuum.

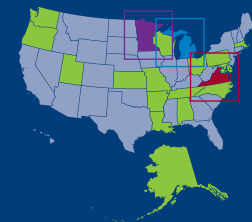
Key Findings for Policymakers

(1) Guidelines make sentences more predictable.

Guidelines substantially achieve their goal of steering courts toward certain sentences for certain types of offenses and offenders. They result in greater consistency in deciding who goes to prison and for how long. Guidelines also produce differentiated punishment: like cases are treated alike while unlike cases result in different degrees of punishment severity. These findings stand in marked contrast to the inconsistent and discriminatory sentencing practices documented in all three states prior to the implementation of guidelines.

More narrow sentence ranges lead to slightly more predictable sentences. Predictability is somewhat higher in Minnesota, where the more mandatory system uses a compact set of sentencing criteria and has relatively narrow sentencing ranges. In contrast, Virginia's voluntary system is based on detailed calculation of sentences but its wider ranges build in more opportunities for the exercise of discretion. Consequently, relatively lower predictability is expected, and found, in Virginia.





(2) Guidelines effectively limit undesirable sentencing disparity.

Guidelines reduce disparities due to factors that should not play a role in sentencing decisions. The undesirable influence of offender characteristics such as race and economic status were of negligible impact in all three states studied.

The discretion afforded judges under more voluntary guidelines does not result in discriminatory sentences. Drawing on the Virginia experience, there is no suggestion in the results of a direct trade-off between predictability and proportionality on one hand and increased discrimination on the other. A voluntary guideline system with substantial sentencing ranges does not necessarily lead to increases in discrimination, as many observers might have expected.

(3) Guidelines make sentencing patterns more transparent.

A valuable by-product of guidelines is that the extent to which they might fall short in achieving predictability, proportionality and non-discrimination is observable and hence correctable through appropriate refinements to the guidelines. There are specific ways that Michigan, Minnesota, and Virginia might improve their guideline policies in terms of redefining their basic guideline elements as well as monitoring sentencing outcomes in their respective jurisdictions. Recommendations for these enhancements are available in a lengthier, companion publication.

(4) State officials have options when designing guidelines.

All guideline systems reflect choices on multiple design considerations about how best to shape judicial discretion. One contribution of the study is the identification of a coherent way to view the similarities and differences in design choices among the many different state systems. The assessment places state guideline systems along a single voluntary to mandatory continuum. This scale allows policymakers to evaluate where their states fit in and to look at other state experiences in tailoring guidelines to match their needs and circumstances.

(5) Active participation by a Sentencing Commission is an essential element of effective guidelines.

Established policies are no more self-sustaining over time than they are self-executing at inception. Sentencing Commissions play a vital role in quality control. They are able to discern if sentences are harmonious with intended goals and make targeted adjustments when necessary. Given the initial purposeful and deliberative investment made by policymakers and commissions to guide sentencing, it is worthwhile to reexamine basic decision-making elements to solidify past and current gains as well as reorient future resources in the most effective manner. Some of the challenges facing the Michigan system might have been avoided through closer monitoring.



Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

Looking Ahead

The evidence and experiences gathered in this examination of sentencing through guidelines should help inform other states considering the introduction of structured sentencing or revisions to existing guidelines. For example, there are critical design considerations and trade-offs related to the appropriate breadth of guideline ranges and the simplicity or complexity of factors used to score convicted offenders.

States continue to examine how best to address the new procedural requirements introduced in the US Supreme Court's *Blakely v. Washington* (2004) and *United States v. Booker* (2005) decisions. Minnesota's sentencing commission has responded to the upward-departure problem identified in *Blakely* by increasing the size of the recommended sentencing ranges. Wider sentencing ranges within the grid cells should significantly lower judicial departure rates, a strategy geared toward making the guidelines "Blakely-proof." A possible unintended consequence is that Minnesota will forfeit a very high degree of predictability – and perhaps proportionality – in this effort to satisfy the strictures of *Blakely*. The results of this study provide policymakers with clear and persuasive empirical evidence of consequences that might follow changes in the guideline structure.

While sentencing guidelines obviously cannot solve every problem and challenge in sentencing and corrections, the study does offer greater clarity on the essential issue of how conscious policy decisions intended to guide judicial discretion affect sentencing outcomes. Future inquiry should explore how alternative sentencing guideline regimes affect the ability of states to effectively manage prison population and control associated costs.

What is the focus of the current research?

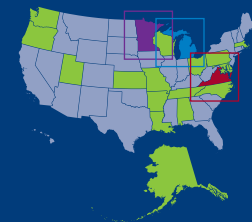
A critical issue is whether the actual sentencing decisions under a guideline framework conform to intended policy objectives. Despite the fact that criminal sentencing has been a perennial topic of analysis and reform for the past several decades, little is known about the character of sentences under guideline systems. In response, the National Center for State Courts has examined and classified all states with sentencing guidelines along a voluntary-mandatory continuum and selected three state systems as representative of alternative ways of configuring the control of judicial discretion (Michigan, Minnesota and Virginia).

Examining the practices in three states, the research asks three questions:

- (1) Are actual sentences predictable using the prescribed elements and mechanics of guideline systems?
- (2) Do more serious offenders receive proportionally greater punishment as prescribed by guidelines?
- (3) Are sentences under the aegis of guidelines fair in the sense of being non-discriminatory, thereby minimizing the effects of extra-legal elements, such as the age, race, gender and geographic location of offenders?

The NCSC develops and applies statistical models designed to simulate the judicial decision-making process by incorporating the information each guideline system provides a judge at the time of sentencing. The models, which consist of statistical equations, are formal representations of the sentencing process. They are tools to estimate and compare what sorts of sentences are predicted (or should be expected to occur) by applying them to information on actual offenders. The models also make it possible to determine whether actual sentences achieve proportionality of punishment along the lines conceived by the guidelines. Finally, the models enable us to address directly the extent to which sentences under these three alternative guideline systems are fair and free from discrimination. Specifically, viewing guidelines in comparative perspective provides insight into understanding how more mandatory guideline systems differ from more voluntary guideline systems.





How do state guideline systems compare?

Drawing on US Supreme Court Associate Justice Louis Brandeis's famous insight, guideline states are "natural laboratories" where sentencing guideline developers have made different policy decisions on their design and operation. The end result has been the creation of sentencing guidelines that take many different forms, despite broad similarities in their intended purpose. Acknowledging the variation that exists among the 21 guideline systems, a coherent way to view them is by comparing them along a common continuum ranging from primarily voluntary recommendations to more mandatory provisions on how judges are to determine appropriate sentences. A direct comparison of states along this continuum makes it possible to examine the impact of alternative design options.

A continuum is created by assigning points to each state based on answers to six questions concerning the state guideline's basic organizational aspects and structural features:

Question 1: Is there an enforceable rule related to guideline use?

Question 2: Is completion of guideline worksheets required?

Question 3: Does a sentencing commission monitor compliance?

Question 4: Are substantial and compelling reasons required for departures?

Question 5: Are written or recorded reasons required for departures?

Question 6: Is appellate review allowed?

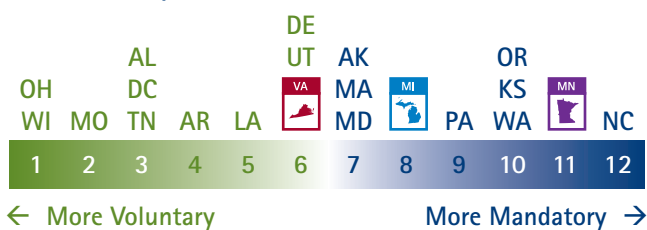
For each question, a state is awarded 0 points for a "no or unlikely" position, 1 point for a "possible or moderate" position, and 2 points for a "yes or likely" position.

Continuum Points Awarded (by Position)

Position	Points
"No or unlikely"	0
"Possible or moderate"	1
"Yes or likely"	2

Summing the points determines the degree to which a state is mandatory or voluntary. States having higher total scores based on all six questions are more mandatory than those with lower scores. The following diagram arrays the states on a single continuum with one pole emphasizing highly voluntary systems (total of one point) and the other pole emphasizing highly mandatory guidelines (total of 12 points).

A Continuum of State Sentencing Guideline Systems






Some states have put in place more mandatory guidelines that more tightly control judicial discretion by using close monitoring, requiring reasons for departures from recommended sentences, and allowing vigorous appellate review. Other states have more voluntary systems where compliance is not monitored, judges are free to depart without having to justify their reasons, and appellate review of guideline sentences is prohibited by statute.



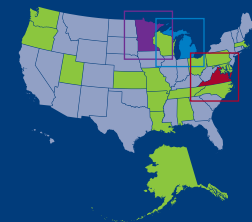
Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

Continuum Questions Applied to Three States: Minnesota, Michigan and Virginia

	Enforceable Rule Related to Guidelines Use?	Worksheet Completion Required?	Sentencing Commission Monitors Guideline Compliance?	Substantial & Compelling Reasons Required for Departure?	Written or Recorded Reasons for Departure Required?	Appellate Review Allowed?
	<p>The Guidelines promulgated by the Sentencing Commission shall establish a "presumptive, fixed sentence for offenders...."</p>	<p>Requires completion of guideline worksheets.</p>	<p>The Commission issues an annual report of guidelines compliance.</p>	<p>Judges are required to give the sentence within the presumptive range. Judges can depart from the presumptive sentence if "there exist identifiable, substantial, and compelling circumstances...."</p>	<p>The judge "must disclose in writing or on the record the particular substantial and compelling circumstances...."</p>	<p>Yes.</p>
	<p>The Michigan guidelines state that, "the minimum sentence imposed by a court of this state...shall be within the appropriate sentence range under the version of those guidelines in effect on the date the crime was committed.</p>	<p>Requires completion of guideline worksheets.</p>	<p>No monitoring of guideline compliance; sentencing commission abolished in 2000.</p>	<p>Judges can "depart from the appropriate sentence range established under the sentencing guidelines... if the court has a substantial and compelling reason for the departure...."</p>	<p>Reasons for departure must be stated on the record.</p>	<p>Yes.</p>
	<p>The Virginia Code specifically states that the guidelines are discretionary.</p>	<p>While compliance with guideline recommendations is voluntary, completion of guideline worksheets is mandatory. Judges are required to review the guidelines in all cases covered by the guidelines and sign the worksheet.</p>	<p>The Commission issues an annual report of guidelines compliance.</p>	<p>Judges are to be given the appropriate sentencing guideline worksheets and should "review and consider the suitability of the applicable discretionary sentencing guidelines..."</p>	<p>In a felony case, if the court "imposes a sentence which is either greater or less than that indicated by the discretionary sentencing guidelines, the court shall file with the record of the case a written explanation of such departure."</p>	<p>No.</p>





Why were Michigan, Minnesota and Virginia chosen for the study?

All sentencing guidelines provide a framework for assessing the severity of criminal activity and a means to arrive at a recommended sentencing range. State guideline systems carry varying levels of authority that circumscribe the discretion of the judge in determining the appropriate sentence. A central issue, then, is how to construct the limits on that discretion and to what end.

To address this issue, three states are selected as representatives of alternative ways of configuring the control of judicial discretion: Minnesota, Michigan, and Virginia. Minnesota is the most mandatory system, followed by Michigan; Virginia is the least mandatory of the three. Minnesota, for example, tends to have tighter ranges on recommended sentences for similarly situated offenders compared to Michigan and Virginia. In addition, Virginia employs a list-style scoring system to determine appropriate offender punishment in contrast to the use of sentencing grids in Minnesota and Michigan. Virginia has one of the most active sentencing commissions, although it is a more voluntary system in terms of requiring compliance, than most states.

What are the critical elements of the sentencing guideline systems in Michigan, Minnesota and Virginia?

The design and operation of the three selected guideline systems are important to describe because their mechanics are incorporated into a statistical model for analysis purposes. Additionally, understanding how the guidelines work in practice is central to examining issues of predictability, proportionality, and fairness.

The mechanics of guidelines involve detailed considerations and calculations, such as how key information on offense seriousness and prior record is handled, how sentences are determined, how sentencing ranges are established, requirements for departures from recommended sentences, whether appellate review is permitted, and how time served is considered.

On the most general level, similarities and differences among the three sets of guidelines are as follows:

Offender Classification

A starting point for the developers of all sentencing guideline systems is how to take into account the interrelationships among:

- (1) **The selection of crime types or crime classifications for inclusion in the guidelines.**

The Michigan grid system distinguishes 9 crime classifications based on statutory severity, the Minnesota grid focuses on 11 offense groups, and Virginia employs worksheets for 15 offense groups.

- (2) **The measurement of prior record.**




Michigan (seven measures) and Minnesota (four measures) use a uniform set of indicators to assess prior record in all cases for all offense categories. Virginia has identified 10 possible prior record variables, but the precise selection, number and scoring varies by offense group.



Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

Structural Comparison of Minnesota, Michigan, and Virginia Sentencing Guidelines Systems

	 MN	 MI	 VA
Commission Status	Active	Abolished	Very Active
Guidelines Format	Single Grid System	9 Grid System	15 Worksheets with Scored Factors
Number of Grid "Cells"	77	258	No cells
Sentencing Range Around Guideline Recommendation	10-15%	50-67%	60-66%
Required Time Served	67%	100%	85%
Aggravated Departures From Recommended Prison Range	29.6%	4%	9.4%
Mitigated Departures From Recommended Prison Range	9%	1.9%	9.4%
Year of Sentencing Data Analyzed	2002	2004	2002

(3) The specifics of the instant offense and is the area where the greatest differences exist between the three systems.

The Michigan guidelines evaluate each offender on up to 20 offense variables, including aggravated use of a weapon, physical and psychological injury to the victim, the intent to kill or to injure, multiple victims, and victim vulnerability among others.

Minnesota incorporates specific offense conduct into the presumptive sentence by imposing mandatory minimum sentences for selected cases involving weapons or second/ subsequent offenses. In Virginia, each offense group has a set of offense conduct variables that apply specifically to that offense (e.g., for Burglary/Dwelling there are six possible aspects of the offense singled out for scoring, such as dwelling occupied, crime occurred at night, intent to use a deadly weapon during the burglary). In addition, there are selected elements of the offense (e.g., weapon type, mandatory firearm conviction) that may apply across many offense groups.

Format

Minnesota and Michigan use a grid system that places offenders into specific cells; Virginia scores each individual offender across a range of variables in a worksheet format.

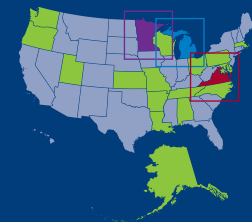
Recommended Ranges for Prison Terms

Michigan and Virginia have wide ranges and base them on past judicial practices. In contrast, Minnesota has narrow ranges based on policy prescriptions concerning what is appropriate and desirable from the point of view of controlling correctional resources.

Permissible Departures from Recommended Ranges

Virginia allows judges to impose sentences that depart from recommended ranges by providing stated reasons, although the sentences are not subject to appellate court review. In Minnesota, judges may depart by disclosing reasons for such action, but the decisions may be reviewed by the Minnesota Court of Appeals. Michigan is similar to Minnesota.





Time Served

In Minnesota, offenders generally serve two thirds of their imposed sentence; in Virginia they serve at least 85 percent. In Michigan, the Parole Board determines the sentence between the judicially imposed minimum, which is served in its entirety (100 percent) and the statutory maximum.

Sentencing guidelines bring together characteristics of the offense and offender in a designed and structured format that weighs or scores an offender and then produces a recommended sentence based on that score. A primary rationale for the choice and weighting of selected factors is to create greater predictability and proportionality and to minimize discrimination in the sentencing process. To date, the relative success of alternative sentencing guideline designs in meeting these fundamental goals remains unresolved.

Why are predictability, proportionality and non-discrimination important when assessing sentencing systems?

Based on organizational structure and process, differences among the three selected state guideline systems are plausibly linked to different sentencing outcomes. Greater understanding of sentencing under guidelines begins with refining the basic vocabulary that describes the characteristics of a desirable sentencing outcome and, by exclusion, a delineation of undesirable outcomes. Clarifying the definition of an acceptable sentence provides a solid base to identify more precisely what are unacceptable deviations. Consistency, for the purposes of this study, focuses on the twin characteristics of predictability and proportionality while fairness focuses on the absence of discrimination.

Predictability in sentencing under guidelines is comprised of two distinct elements.

- (1) Sentences are predictable to the extent similar offenders receive similar sentences.
- (2) Sentences are predictable to the degree individual offenders are placed into distinctive groups, each with a range of justified punishment based on a "legitimate" set of characteristics.

Sentences should also be proportional; that is, dissimilar offenders receive dissimilar sentences in proportion to their degree of dissimilarity.

Under guidelines, the goal is to make sure more blameworthy crimes receive more severe punishments.

A primary task of sentencing guideline designers is to make concepts like "similarly situated," "range of justified punishment," and "more blameworthy" precise and measurable. For example, a given combination of offense seriousness and prior record on the Minnesota guideline grid locates and defines a set of offenders deemed to be similarly situated. Being in the same grid cell carries the implicit prediction that the offenders are of comparable blameworthiness and hence should receive similar penalties.



Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

Likewise, successive steps up or down the offense seriousness and prior record scales identify dissimilar offenders as well as the extent to which they are dissimilar. In Minnesota, for example, if two offenders are convicted of the same offense, the offender with a higher level of prior record score will be recommended for a more serious sentence. Guidelines define a series of thresholds that represent jumps from one level of blameworthiness to another. Because crossing a threshold carries an increase in the severity of penalty, it is important that adjacent levels should be formally and meaningfully distinct from one another. If not, proportionality is violated.

On the most general level, discrimination refers to sentences that are different, with the source of the difference tied to specific extra-legal characteristics of the defendants. For this reason, the current research focuses on the kinds of undesirable disparities guidelines are designed to prevent – those resulting from the offender’s race, age, gender, the region of the state in which an offender is sentenced (the key question with regional variation is whether there are distinct sentencing “regimes” operating under the banner of a single sentencing guidelines structure), and the manner of disposition. Minimizing the effects of these sources of potential discrimination is an explicit goal in all three systems examined.

In summary, three criteria related to predictability, proportionality, and discrimination guide the current evaluation of whether more voluntary guidelines perform differently than more presumptive ones.

- (1) Do similarly situated offenders as defined by the guidelines receive similar sentences?
- (2) Do the guidelines provide meaningful and proportional distinctions between more serious and less serious offenders?
- (3) Is there evidence of discrimination in sentencing?

What methodology was used to evaluate consistency, proportionality, and discrimination in sentencing?

From a research perspective, the legal policy outcomes (or dependent variables in statistical models) to be explained correspond to the following two types of sentencing decisions:

(1) Who is sentenced to prison?

The decision is whether to punish a defendant convicted of a felony offense with a prison sentence or to impose a less severe penalty, typically involving some combination of jail, probation, fines, work release, therapeutic treatment, and restitution. The choice between these alternatives is commonly known as the “in/out” decision.

(2) What determines the length of time an offender is sentenced to prison?

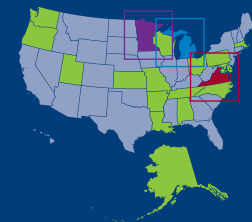
Aptly characterized as the prison length decision, the analysis focuses on identifying the factors influential in determining sentence duration.

A broad range of factors related to offense and offender characteristics (the independent variables in statistical models) are included to determine how they affect sentencing outcomes:

Essential elements and mechanics of each guideline system

These variables are tailored to fit the unique features of each guideline system and generally include measures of the offense at conviction, prior criminal history, specific conduct surrounding the offense, the type of grid (Michigan and Minnesota) or guidelines score (Virginia), the offender’s habitual offender (Michigan) or modifier (Minnesota) status, and any guidelines departure (if applicable) from the recommended range.





Extra-legal factors

This set of variables includes measures on an offender's age, race, gender, and geographic region of the state. Statistical models were constructed to estimate (or predict) the two sentencing decisions (i.e., whether sentenced to prison and if so, for how long) for each offender in each state based on information on offenders' characteristics and the elements and mechanics of each system. Estimates can be made whether the information and guideline elements call for imprisonment and if so, for how long for every individual offender. The information on offenders was drawn from a large number of cases in each state; Michigan (N=32,754), Minnesota (N=12,978) and Virginia (Assault N= 1,614; Burglary N=1,668). A comparison of the actual sentencing decisions to the predicted decisions when the statistical models are applied to information on offenders convicted of felony offenses reveals how successful the sentencing guidelines are in terms of achieving predictability, proportionality and non-discrimination.

Predictability Tests

The specific criteria used by the guideline designers to define the concept of similarly situated are used to evaluate the internal workings of each guideline system. The analysis examines whether sentence outcomes follow in a predictable manner from the combination of offense and offender characteristics built into the guideline system. Are offenders sentenced on the basis of the set of elements provided for in the guidelines? In statistical terms, do the sentencing guideline factors account for the observed variation on sentencing?

Proportionality Tests

The focus here is whether the individual guideline elements related to offense severity and prior record have a proportional impact on sentencing. For example, there are six levels of offense seriousness in the Michigan guidelines: Is the impact of each level distinct – in a statistical sense – from the adjacent levels? A movement between levels carries direct consequences for convicted offenders in terms of exposure to prison time. Because guideline designers elected to make these distinctions, whether the individual intended differences in sentencing outcomes correspond to actual judicial choice is an empirical question. Therefore, it is anticipated that more serious classifications of offense and prior record will be associated with higher estimated probabilities of receiving a prison sentence and longer prison sentences.

Discrimination Tests

By examining the statistical coefficients associated with the impact of each of the extra-guideline variables, the extent to which a system minimizes discrimination in sentencing is discernible in measurement terms. The potential influence of age, gender, race and their interactions with each other (e.g., young, black men) and other variables (e.g., state geographic regions) are examined in considerable detail to determine if guidelines are sufficiently successful in promoting predictability and proportionality to the point that discrimination is minimized.



Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

What did the study reveal?

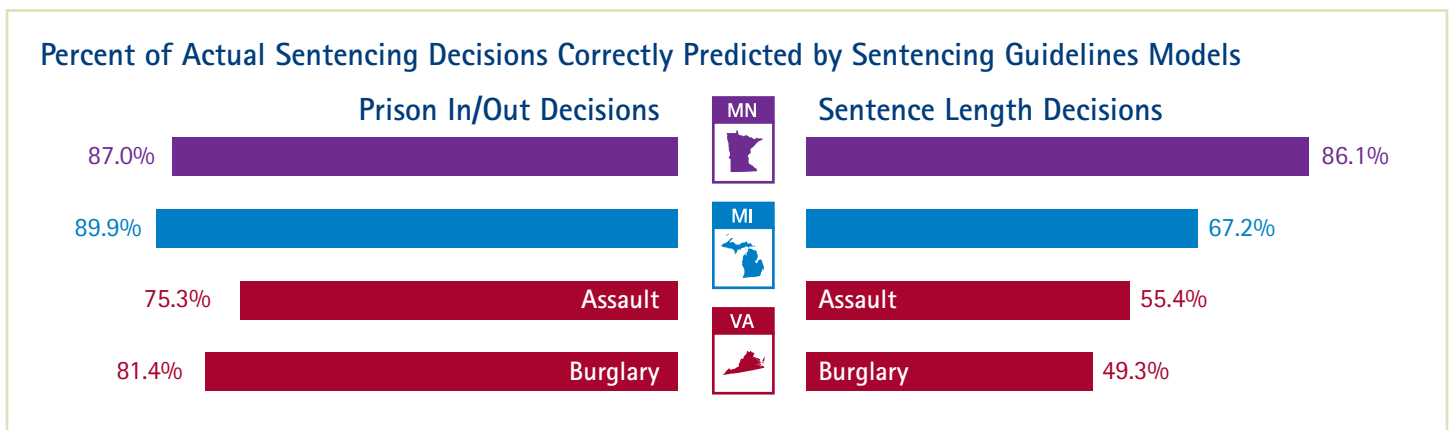
Predictability

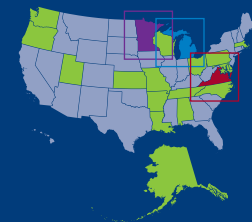
Do actual sentences correspond to sentences suggested by guideline criteria and mechanics? The evidence indicates a close overall fit between predictions based on the guideline elements and reality. A model of the In/Out decision in Michigan predicts 89.9 percent of the cases correctly; the Minnesota model predicts 87 percent of the cases correctly; and the Virginia model correctly predicts 75 percent of Assault offenses and 81 percent of burglary offenses. Hence, despite differences among Michigan, Minnesota, and Virginia in guideline design and structure, the three sets of guidelines work effectively to guide judges in a predictable manner in making the basic in/out decision.

Predictability also refers to how well an offender's placement on the guideline grid (or worksheet score in Virginia) relates to the actual length of prison sentence received. Looking at the full range of prison sentences received by convicted offenders in a particular state

shows a great deal of variation from relatively short, say 1 year, to very long, say, 50 years. If the guidelines are operating as envisioned, most of the variation in sentence length should be related to differences in the specific offense and offender characteristics scored as part of the guideline calculation. A key question, then, is what proportion of variance in observed sentence length is explained by the guideline factors?

For Minnesota, the statistical model accounts for 86 percent of the variation in sentence length followed by Michigan (67 percent) and Virginia (53 percent for Assault and 49 percent for Burglary offenders). While the proportion of explained variation is related to where the system is on the voluntary/mandatory dimension, the predictability in sentence length is substantial in all three guideline systems. Taking both the in/out and sentence length decisions together, all three guidelines have dramatically enhanced the predictability of sentencing.





Proportionality

A second key aspect of consistency under guidelines is that similarly situated offenders receive similar sentences. Conversely, dissimilar offenders should be treated differently. Proportionality is a value that functions as a principle in determining what “different” means. Simply stated, proportionality entails a balance between the severity of the offense and prior record and the degree of punishment.

While the full report examines a series of refined tests of the degree to which proportionality exists, a look at two types of tests in the context of the in/out decision provides insight into the issue. The first test asks: is there a statistically significant difference between the likelihood of offenders being sentenced to prison who are in different locations on the Minnesota and Michigan sentencing grids? For example, do judges make significant distinctions between adjacent prior record levels in the Michigan guidelines when imposing sentences? If so, this information indicates the formal levels built into the guidelines are efficacious in drawing distinctions between similar and dissimilar offenders. Such analysis helps address whether judges in their actual sentencing decisions employ proportionality when making a horizontal or vertical move between grid cells.

Policymakers institutionalize jumps in the recommended severity of punishment following changes in discrete offense or prior record thresholds. Examining the column labeled “Percent change in probability of going to prison” on the table above shows how an increase in the seriousness of prior record or the seriousness of the offense changes the estimated probability of receiving a prison sentence and whether the change is statistically significant.

Assessing Proportionality in Michigan and Minnesota (Seriousness level increases from low to high for each variable listed)



Michigan In/Out Decision

Variable	Level	Percent change in probability of going to prison
Prior Record (comparing % change from Level A)	Level B	-2%
	Level C	2%
	Level D	12%
	Level E	28%
	Level F	38%
	Offense Seriousness (comparing % change from Level I)	Level II
	Level III	8%
	Level IV	10%
	Level V	13%
	Level VI	16%



Minnesota In/Out Decision

Variable	Points/Level	Percent change in probability of going to prison
Criminal History (comparing % change from 0 points)	1 Point	15%
	2 Points	32%
	3 Points	36%
	4 Points	45%
	5 Points	53%
	6 Points	67%
Severity of Conviction Offense (comparing % change from Level II)	Level I	3%
	Level III	3%
	Level IV	6%
	Level V	10%
	Level VI	29%
	Level VII	67%
	Level VIII	53%
	Level IX	84%
	Level X	84%
	Level XI	84%

Note: All changes in probability of going to prison were statistically significant except for the change from Level I to Level II for Minnesota Severity of Conviction Offense.



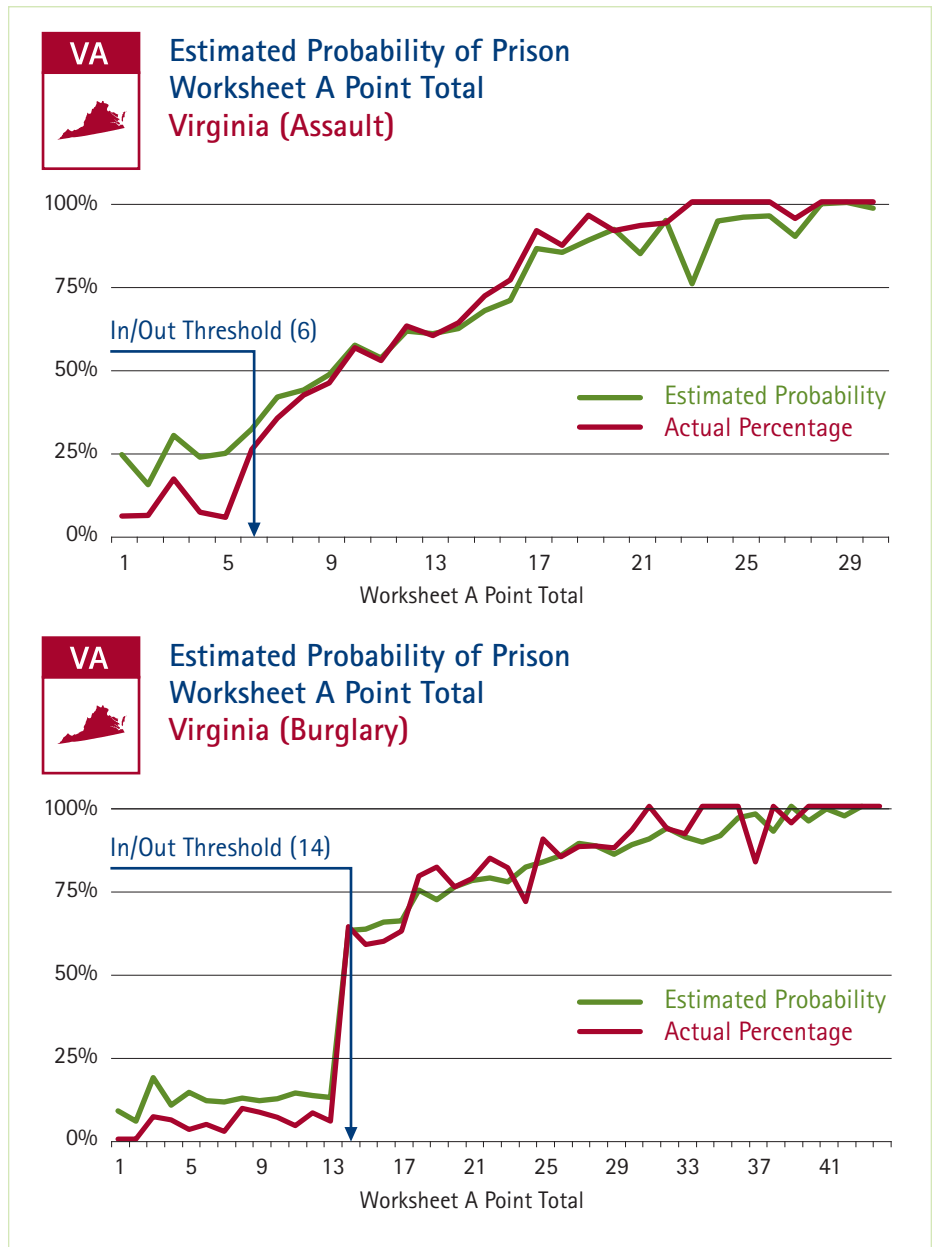
Assessing Consistency and Fairness in Sentencing:

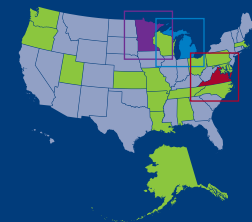
A Comparative Study in Three States

The probabilities can be interpreted as the change in likelihood of going to prison for offenders found to be more serious than a lower level baseline offender. For example, in Minnesota, the baseline offender is an individual with 0 criminal history points. An offender who is similar in all respects to the baseline offender except with a criminal history score of 4 points has a 45 percent increase in the likelihood of receiving a prison sentence.

Hence, a change in probability should increase with an increase in the level of seriousness. More serious offenders, as measured by more extensive prior record or more serious conviction offense, should have a higher probability of prison. Indeed, this is what is found. Almost all distinctions are statistically significant (the exception is no statistical difference between offense severity Levels I and II in Minnesota) and in the right direction (the exception is that prior record level B in Michigan is found to be statistically significant in the opposite direction expected). Consequently, the guidelines demonstrate effectiveness in distinguishing more serious from less serious offenders and in leading judges to sentence offenders accordingly.

The Virginia guidelines are used to illustrate a second approach to assessing proportionality. Whether the guidelines recommend an offender receive a prison sentence is determined by scoring a range of offense and offender characteristics, totaling the points, and comparing this total against





an established threshold value. If the offender's score exceeds the threshold, the guidelines recommend a prison sentence. The total score provides a judge with an immediate summary assessment of each offender that is directly comparable to the threshold value at the bottom of the worksheet. Moreover, higher scores indicate proportionally more serious offenders in the context of the Virginia guidelines. The concept of proportionality implies that as the total score increases, there should also be an increase in the likelihood of prison.

The results show offenders in Virginia with lower total worksheet scores are less likely to receive a prison sentence than offenders with higher scores. Both the Assault and Burglary figures provide strong evidence of proportionality, however, there are differences between the crime groups as shown in the adjacent figures. The figures present both the actual percentage as well as the estimated probability of prison for offenders of varying seriousness, according to the guidelines, for Assault and Burglary. For Assault, the predicted probability of prison is only 30 percent at the threshold value of 6 points and does not reach 50 percent until a total of 10 points is reached. For the Assault crime group, the judges appear to exercise discretion, as is their right under a voluntary system, in determining whom to incarcerate. In practice, the threshold acts more as a strong signal than a strict legal standard.

For Burglary, the figure shows that below the threshold of 14 points the probability of receiving a prison sentence is stable at a very low rate. However, once the point total exceeds 13, there is not only a dramatic jump in the probability of prison but the probability continues to rise as the worksheet total increases. In this case, the threshold is operating as envisioned by the guideline designers and creates a sharp discontinuity when the total score exceeds the threshold value. The results indicate judges are following the overall guideline recommendation for the in/out decision.

Discrimination

A critical aspect of the NCSC research is to examine the extent to which any observed inconsistency in sentencing can also be called discriminatory. Discrimination is a particularly troubling type of inconsistency as it implies offenders are treated differently based largely on morally and legally undesirable criteria. A potential confounding factor is that sentencing outcomes may vary by region around a state. One implication of the "similarly situated" concept under statewide guidelines is that similarly situated offenders are treated similarly in all parts of the state. Therefore, geographic variation is also examined as a source of unwarranted disparity.

The results reported here come from a battery of refined statistical tests. In discussing and evaluating them, a critical distinction between statistical versus substantive significance should be underscored. Sensitivity to this difference is warranted especially with controversial topics like sentencing discrimination. Just because a factor is found to be *statistically* significant does not mean the effect is *substantively* significant; that is, that it has a large effect on the outcome. A variable might be statistically significant but have a very small impact that does not reflect substantial differences in the real world.

The news from the current research is that while a small number of statistically significant racial effects were found across the three states, all were substantively small with minimal impact on actual sentence decisions. For example, while race alone is not significant in Michigan and Minnesota, the subgroup of young black males has a slightly greater chance of being sent to prison of less than one percent. In Virginia, the guidelines have eliminated almost all evidence of racial differences in sentencing across the six crime groups examined with one exception. Black males register a slight increase in predicted sentence length for the Assault crime group.



Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

With respect to males and females, there are statistically significant findings across all three guideline systems that female offenders are treated more leniently both with respect to the in/out decision as well as the prison length decision. However, the substantive impact of these differences is typically small. For example, all other things equal, women have less than a one percent lower probability of being sentenced to prison in all three states.

Michigan is the only system where age was found to have an impact. Older offenders are marginally more likely to go to prison. However, even in this state, age was not found to affect the length of sentence.

While there is little evidence of direct discrimination due to race, age, or sex, the analysis suggests that there is a less obvious source of discrimination brought on by the differences in sentencing outcomes between the large urban courts and the rest of the state – especially in Michigan. To varying degrees, the operation of local norms can sometimes circumvent the goal of statewide uniformity in sentencing. And there is evidence that the informal rules and norms in the large urban courts shaping what sentences are deemed appropriate differ from courts in the rest of the state. While the analysis shows that the differences are statistically significant, it is clear that, at least in Michigan, the differences are substantively significant as well.

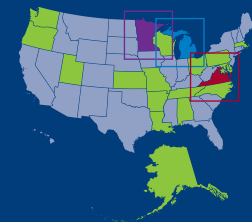
Offenders in metropolitan Southeast Michigan (which include 60 percent of all black offenders) receive sentences that are markedly more lenient than their counterparts in the rest of the State (or out-state). Results indicate the probability of going to prison is 10-15 percent higher in out-state Michigan and the length of sentence is 25-30 percent greater. A single set of guidelines is being applied in a very different manner in different parts of the state.

The analysis suggests the primary reason for the presence of two statistically and substantively significant sentencing regimes in Michigan can be traced to the very large guideline sentencing ranges. The magnitude of the ranges means that judges can sentence quite differently without having to depart. If the norms of the urban courts lead judges to look to the bottom of the ranges, while out-state judges look toward the top, there can be dramatic differences in sentencing outcomes. While there is little evidence of discrimination as usually conceived, geographical disparities undermine the goal of statewide consistency.

In Minnesota the geographical differences are smaller and different than in Michigan. Hennepin County, the state's most populous county, has a slightly higher rate of imprisonment and slightly shorter sentences. In order to mete out shorter sentences within the confines of a guideline system with very narrow ranges, it is not surprising that Hennepin judges depart below the recommended guideline range twice as often as do judges in the rest of the state.

One line of thought suggests that since the Virginia sentencing guidelines are voluntary, there is more room for judges across the Commonwealth to treat convicted offenders differently. However there is no evidence to suggest that there is systematic discrimination – that rises to the level of statistical significance – in Virginia. This is interesting given that the explained variance in both Virginia crime groups is less than that of the Michigan and Minnesota counterparts. With more variation unexplained, it seems likely to find some systematic discrimination, however no supporting evidence for this was found in the current research.





What conclusions can be drawn from the study?

From the enquiry into the application of sentencing guidelines in the three states of Michigan, Minnesota and Virginia, there are five broad conclusions that increase the understanding of how sentencing guidelines work to shape and control the discretion of trial court judges.

The main conclusions are:

- (1) Guideline systems produce predictable sentencing decisions based upon their prescribed elements and mechanics.

In addition, the guidelines result in differentiated punishment. Like cases are treated alike while unlike cases result in different degrees of punishment severity. Finally, the undesirable influence of extra-legal factors is negligible in all three states.

- (2) Predictability is somewhat higher in the context of Minnesota's more compact set of elements and use of relatively narrow guideline ranges.

However, with the compactness comes a higher propensity for departures. In contrast, Virginia's more detailed system allows for greater flexibility in how the guidelines are to be applied (i.e., more voluntary), thus building in more opportunities for the exercise of appropriate discretion.

- (3) There is no evidence of a direct trade-off between predictability and proportionality on one hand and undesirable racial, gender, or age disparities on the other.

In fact, a voluntary guideline system, such as the one in Virginia, with substantial sentencing ranges exhibits no measurable discrimination.

- (4) All guideline systems benefit from periodic assessment of current practice and the extent to which the guideline systems are achieving key goals of consistency and fairness.

Information on actual practices provides clear and interpretable grounds for adjusting guideline elements and mechanics. As a result, increased accountability in future sentencing can be promoted on the basis of past performance and not just on the basis of conjecture or supposition.

- (5) Finally, policymakers, judges and all others concerned about sentencing will benefit from working together to ensure the establishment of sentencing commissions to regularly monitor sentencing patterns to solidify past and current gains as well as reorient future resources in the most effective manner.

These conclusions underscore the value of comparative research in criminal sentencing by clarifying the similarities and differences in sentencing guideline structures and their respective patterns of outcomes. Only comparative enquiry provides an understanding of where the differences lie and what their consequences might be. Hence, it is hoped future researchers continue to probe the conduct and outcomes of sentencing across states and develop a broad base of data on which conclusions are reached. More comparative enquiry will help inform policymakers in their deliberations on sentencing guidelines.





Assessing Consistency and Fairness in Sentencing:

A Comparative Study in Three States

The National Center for State Courts is an independent, nonprofit, tax-exempt organization in accordance with Section 501 (c)(3) of the Internal Revenue code.

The National Center for State Courts is headquartered in Williamsburg, VA, with offices in Denver, CO, and Arlington, VA.



NCSC
300 Newport Ave.
Williamsburg, VA 23185-4147
www.ncsconline.org