

Sentencing Using Statistical Treatment Rules: What We Don't Know Can Hurt Us

Shawn Bushway
School of Criminal Justice
State University of New York at Albany
SBushway@uamail.albany.edu

Jeffrey Smith
Department of Economics
University of Michigan
econjeff@umich.edu

April 21, 2007

We thank Peter Reuter and Gary Sweeten for helpful comments and discussion and three anonymous referees for their thoughtful comments. Despite our best efforts to place the blame elsewhere, all errors remain our own.

“Sentencing Using Statistical Treatment Rules: What We Don’t Know Can Hurt Us”

Abstract

This paper makes the case that the existing literature seriously misinterprets the available evidence on the predictability of high risk criminal and thus on the potential value of statistical treatment rules aimed at providing stiffer punishments to offenders with higher predicted risk of recidivism. The misinterpretation results from the failure to take account of the fact that the data used in existing risk assessment exercises come from environments characterized by informal (and sometimes formal) attempts by judges and other actors to base penal treatments on expected recidivism. As a result, findings of little predictive power for baseline covariates may simply indicate the efficient use of the available information. We lay out the problem in detail, provide examples from several literatures and then consider general solutions to the problem.

Keywords: selective incarceration, sentencing, statistical treatment rule, profiling

Running head: Sentencing Using Statistical Treatment Rules

I. Introduction

Perhaps the best known fact in criminology is that a small percentage of offenders are responsible for the majority of all crime. Piquero and Blumstein (this volume) note that this strong positive skew raises numerous problems for researchers trying to estimate the “average” impact of incapacitation. At the same time, this reality raises the provocative possibility of decreasing crime by selectively incapacitating these high rate offenders, if we could identify them prospectively.

Over several decades, criminologists have devoted their intellectual efforts to the problem of identifying high-rate offenders using prison samples; see the review in Harcourt (2006) of the long history of risk prediction in criminology. As noted by Gendreau et al. (1996), these efforts have focused primarily on criminal history records as predictors of future behavior. The general consensus in the criminological literature – see e.g. Gottfredson and Gottfredson (1994) and Auerhahn (1999) – views these efforts to identify high risk offenders as largely ineffective.¹

But not all practitioners of risk assessment have been as deferential to this prevailing logic. Research in this area has accelerated as demonstrated by a recent special issue of *Crime and Delinquency* devoted to the topic, and it is fair to say that this new wave of risk prediction represents a renewed effort at selective incapacitation. Indeed, more than half of the U.S. states now use some type of risk assessment to identify and detain high risk offenders as part of their parole system. Of course, this movement back towards selective incapacitation has not gone unnoticed by the major critics of the last wave of selective incapacitation research. For

¹ At the same time, a different consensus has emerged regarding the ability to predict behavior at the other end of the offender distribution. Visher (1986) and Gottfredson and Gottfredson (1986), among others, demonstrate the feasibility of prospectively identifying low rate offenders with reasonable accuracy. Some practical applications of statistical treatment rules have followed these disparate conclusions in the literature by focusing on identifying and treating less intensively individuals predicted to have low offense rates.

example, Gottfredson and Moriarty (2006) argue that this new round of research ignores the lessons of earlier research and therefore creates invalid instruments. They claim not that risk prediction cannot be done, but rather that the methods employed in recent analyses have so many methodological problems as to make the exercises largely uninformative. They hold up earlier work done in criminology, such as the oft-cited paper by Gottfredson and Gottfredson (1994) that uses 25 years of follow-up data on California prisoners, as exemplifying the proper way to create and validate these risk prediction scales.

In this note, we go beyond the concerns of Gottfredson and Moriarty (2006), with whose critique we generally agree, and identify a serious conceptual problem that can be found even in seminal papers such as Gottfredson and Gottfredson (1986, 1994). The problem arises from the indisputable fact the data used in the formal risk assessments in the criminal justice and criminology literature are generated by processes of informal risk assessment and treatment assignment. Put somewhat differently, the available data embody efforts by judges, prosecutors, parole boards and other actors in the criminal justice system to use the information available to them (which includes information observed and not observed by researchers) to predict risk and to assign punishments based on those predictions. If these treatments work, then the behavior we observe for the high risk people will in fact be less risky than it would have been otherwise. Berk and de Leeuw (1999) and Berk et al. (2003) refer to this problem as the “suppressor effect” of treatment. As a result, it is impossible without additional strong assumptions to distinguish the “true” behavior of individuals from the behavior that results from their non-random treatment within the existing system. Failure to “undo” the effects of non-random assignment to penal

treatments under the existing regime in empirical risk assessments will lead to misleading estimates and misled researchers.²

The remainder of the paper proceeds as follows: Section II considers the fundamental problem with identifying high rate offenders so as to subject them to harsher punishments via a statistical treatment rule. Section III considers ways to solve the problem and, as a byproduct, makes it clear that doing so requires advances beyond the existing literature in both data and methods. Section IV concludes.

II. The problem and its implications for interpreting evidence on prediction

This section considers four examples of the basic problem that determining our ability to predict high rate offenders, a necessary prerequisite to differentially treating predicted high rate offenders under a statistical treatment rule, depends on undoing the treatment rule implicit in existing data. These examples illustrate clearly that failure to undo the existing treatment regime can lead researchers to completely misunderstand the nature of the empirical evidence they produce. The general problem considered here arises in a variety of different disciplines and substantive contexts; we mention some of these other contexts in this section and the following section because they illustrate the generality of the issue and because they highlight potentially interesting solutions not well illustrated in the criminological literature.

² Our argument parallels a similar (and independently developed) argument in Kuziemko's (2006) thoughtful study of parole, though we frame our discussion at a higher level of generality. The argument can also be seen as an extension of Berk and de Leeuw's (1999) important insight regarding the suppression effect of treatment. See also Berger, Black and Smith (2001) and the references therein (particularly the papers by Charles Manski) for general discussions of the economics and econometrics of statistical treatment rules.

A. WPRS

Consider first an example from outside the criminological context, namely the Worker Profiling and Reemployment Services system (WPRS) analyzed in Black et al. (2003) and Black et al. (2007). The WPRS system assigns mandatory reemployment services to Unemployment Insurance (UI) claimants using a statistical model that predicts claim duration; each state constructs and estimates its own model. In general, the law entitles new UI claimants to 26 weeks of benefits. Claimants predicted to collect most or all of these benefits get assigned to the mandatory treatment and the others do not. The statistical models for this system rely on data prior to the introduction of WPRS to predict claim duration in the absence of WPRS.

Suppose now that a particular state wants to update their statistical model. Suppose further that the state thinks the old data do not reflect current labor market realities. If the state wants to update its model it must do so using data generated during the WPRS era. That is, it must estimate the probability of having a long claim in the absence of WPRS using data generated in the presence of WPRS.

To see why this might be problematic, further suppose that the claimant population has two observable types of individuals, those whose spells in the absence of WPRS would last four weeks and those whose spells would last six weeks. Suppose further that the current WPRS system in the state assigns a treatment to the second type of individuals that reduces their spell duration by two weeks. The end result is a sample of individuals all of whom have a four week spell of unemployment. A simple regression of claim length on observed type yields a coefficient of zero on the type variable and an r-squared value of zero. Of course, estimating such a model provides no information about the potential value of having a statistical treatment rule, which could be quite large if the mandatory services have a smaller effect on the first type

of claimants than the second. This regression also does not show us how to construct a useful treatment rule.

Yet this meaningless regression in our simple WPRS example represents the exact analogue of the many regressions of arrests or recidivism on covariates estimated in the criminology literature. These meaningless regressions then get used as evidence for the proposition that statistical treatment rules have no potential value in the sentencing context because it is impossible to predict high-risk offenders.

B. Racial profiling by police

The recent introduction of outcome analysis into the discussion about racial profiling by police provides a second straightforward example of our general point. Knowles et al. (2001) argue that researchers should examine the hit rates of the searches, where a “hit” occurs when officers find something, such as a gun or some illegal drugs, rather than the racial makeup of traffic stops. They formally demonstrate that in a world where troopers seek only to maximize hits, where the observable characteristics of drivers provide information about the probability of a hit, and where troopers use this information efficiently when making stops, hit rates from observed stops will be uncorrelated with the observable characteristics. In such a world, a risk prediction exercise using data from actual stops would yield an r-squared value of zero.

As in the WPRS example, in this example the fact that a risk prediction exercise using the observed data would turn up nothing does not mean that risk prediction does not work; indeed, in this case it means that existing practice makes use of all of the information in the observable characteristics. Furthermore, a non-zero coefficient on a particular variable in the risk prediction exercise means that troopers can do better (i.e. increase the number of hits for a fixed number of

searches) by either searching more (a positive coefficient) or fewer (a negative coefficient) cars with drivers having that characteristic.

Of course, these conclusions depend on the correctness of the model, in particular the assumption that state troopers seek to maximize the hit rate and the hit rate alone when choosing which cars to search. As Knowles et al. (2001) acknowledge, this interpretation is invalid if the troopers are not maximizing hit rates (see also Engel forthcoming). Our interpretation of a risk prediction exercise depends critically on our understanding of the goals of the actors (in this case state troopers) in the current system. Similarly, in the sentencing context, the interpretation of a risk prediction exercise like the one in Gottfredson and Gottfredson (1994) depends critically on what the researcher thinks the judges and other actors (and the sentencing rules and guidelines) seek to accomplish.

C. Parole in Georgia

Kuziemko (2006) analyzes the behavior of parole boards in Georgia and provides two pieces of evidence consistent with our main point. Her most striking piece of evidence makes use of a court-ordered prisoner release that superceded existing parole timing decisions. As a result of this unique policy shock, she observes a sample of prisoners for whom she knows the parole board recommendation but for whom that recommendation did not get implemented. An analysis of recidivism using the mass release sample indicates that the effect of a one month longer recommended sentence more or less equals her estimate of the effect of an additional month of incarceration on recidivism. On the basis of this finding, she infers that "... parole boards try to assign prison time so that all inmates exit with roughly the same expected recidivism rate."

She then tested this idea with a separate analysis of a random sample of prisoners for whom she has same information relied on by the parole boards in their decisions. The analysis makes use of the fact that, in the period covered by her data, parole boards have the power to locate the realized sentence between the time of the initial review (about a quarter to a third of the way into the sentence) and the judge's original sentence, but have no power to move the realized sentence outside that range. As such, if parole boards use the information contained in the observable characteristics to release offenders when they hit a certain level of risk, these variables will have no predictive power for those over whom the parole board has relevant discretion. On the other hand, the variables should continue to have predictive power for prisoners released shortly after the initial review (for whom the observables may have suggested a shorter sentence than the parole board could impose) or released at or close to the end of the original sentence (for whom the observables may have suggested a longer sentence). Kuziemko's data reveal exactly this pattern.

In her example, as in the others, lack of explanatory power indicates the success, not the failure, of risk assessment conditional on observables under the existing regime. Thereby, it similarly suggests the potential value of basing a formal statistical treatment rule system on such predictions. More generally, the standard conclusion that poor fit in a behavior prediction equation implies that the behavior is not in fact predictable depends critically on the data generating process. If the data come from an environment with random treatment assignment, poor fit does indeed indicate lack of predictive power. In contrast, if the data come from an environment of purposive treatment assignment, poor fit may not indicate an underlying lack of predictability, depending on the nature of the existing assignment regime.

D. Virginia Criminal Sentencing Commission

A final example comes from the Virginia Sentencing Commission. The analysis sample employed to develop the profiling system revealed age as a substantively and statistically significant predictor of recidivism. As a result, age became a major feature of the profiling scheme. In contrast, in data collected after the implementation of the profiling system (with its concomitant assignment of treatment based in large part on age) age no longer predicts recidivism. Ostrom et al. (2002) interpret this finding as indicating that age lacks the predictive power suggested by the initial analysis. In a more recent paper, the same authors now suggest that this means that the information in age is being used efficiently to minimize crime (Kleiman et al., 2007). The interpretation of the age variable in this case depends critically on comparing the models under two known treatment regimes.³

E. Interpreting the examples

In each example, a treatment regime already in place influences the outcomes observed in the data. As a result, in these examples the standard prediction exercises in the literature do not have the simple and obvious (but incorrect) meaning that the literature assigns them. Instead, a finding that observable characteristics do not predict unemployment duration, or recidivism, or hit rates in auto searches in the observed data conditional on the existing treatment regime means only that the existing regime effectively makes use of the information in these variables.

Of course, just because it follows that if the current system efficiently uses the available observable characteristics in assigning treatments standard risk assessment exercises will suggest a lack of predictive power does not mean that the same relationship necessarily holds in the other

³ Note that the analysis used to create the tool relied on a sample of released prisoners, while the validation analysis relied on a sample of deferred probationers. This difference in samples could also explain the result.

direction. In any given context, a finding of no predictive power might mean exactly that – it all depends on the nature of the existing treatment rule. In addition, a finding that a given set of observable characteristics have no explanatory power does not mean that other characteristics, observed by actors in the system but not by the researcher, do not. In the next section, we discuss potential solutions to this problem.

III. Potential solutions

Existing treatment decisions at all stages of the penal process depend in part, and sometimes in large part, on the discretionary actions of individuals such as judges, probation officers and so on. Undoing the statistical treatment rule implicit in these actions in existing data or in data collected in the future so as to allow the consistent estimation of the effects of alternative rules represents a major intellectual task. Progress on this front depends, first and foremost, on the literature coming to understand the largely unappreciated point made in this paper. It will also require developments in statistical methods and data collection, as well as institutional flexibility and change. At least one anonymous reviewer raised serious questions about how realistic such developments would be and rightfully criticized us for doing nothing more than pointing out the problem. We understand the pessimism, but think a brief discussion of the necessary steps is illustrative at least. Therefore, this section briefly considers, at a very broad level, the alternative approaches to obtaining this information.

The problem of undoing the existing implicit treatment rule represents an especially difficult task in many criminal justice contexts for several reasons. First, the time horizon required to fully observe the outcomes associated with alternative treatment regimes often

consists of years, rather than months or weeks, as in other contexts such as UI profiling. This complicates data collection and slows the pace of knowledge creation.

Second, the existing criminal justice treatment regime defies concise description or modeling. Current practice is largely clinical in nature with substantial discretion, and we have little detail about how treatments get selected. In addition, the sentencing literature provides substantial evidence of multiple goals of sentencing (Steffensmeier 1980; Steffensmeier et al., 1993 and 1995; Tonry, 1996) whose relative importance varies across actors (Gottfredson 1999). None of the risk assessment instruments currently in the literature attempt to model the treatment rules implicit in the current regime. Without an accurate model of the treatment regime, we simply cannot assess the true predictive capacity of any given variable under an alternative treatment regime.

Third, and perhaps most importantly, we have what Kuziemko (2006) usefully labels as the simultaneity problem and Berk and de Leeuw (1999) calls the suppressor effect. As we have repeated throughout the paper, the interpretation of the evidence from an empirical risk assessment exercise depends on the motives and actions of the actors under the treatment regime implicit in the data. Actors in the criminal justice system act upon the convicted offenders by assigning treatment, treatments that can then affect recidivism both directly, via incapacitation or specific deterrence, or indirectly via general deterrence effects. Put differently, treatment and recidivism risk are jointly determined. Recidivism risk affects treatment, and treatment can affect recidivism risk. The variables commonly used in risk prediction exercises also drive treatment, and we have no simple way to disentangle the true impact of treatment on recidivism, or, more importantly, to obtain proper estimates of the impact of the prediction variables on recidivism risk. For example, prior criminal history increases risk, but it also induces treatment

(incarceration) that could alter behavior and our ability to observe that behavior. This clouds the meaning of variables measuring the effects of past criminal history on recidivism in exercises like that in Gottfredson and Gottfredson (1994). For the same reason, the current risk assessment tools reviewed in Andrews et al. (2006) do not provide clear answers to the question they address because they fail to take account of this simultaneity.

Consider first two solutions likely unavailable in the sentencing context. First, in the WPRS context discussed in the preceding section, researchers initially had available data uncontaminated by the WPRS or any other treatment regime because prior to WPRS UI claimants never had to participate in reemployment services (other than the trivial one of signing up at the Employment Service) in order to keep their benefits. Even now, a state that really wanted to re-estimate its profiling model on more recent data could simply turn off the WPRS system for a few months and collect all the information it needed. In contrast, we cannot even imagine “turning off” the criminal justice system in order to observe clean data on recidivism and, even if we did, because criminal activity reacts to expected punishment, the data would not provide the desired information as they would reflect criminal behavior with zero punishment, which constitutes an uninteresting scenario from a policy point of view.

Second, in the literature on statistical treatment rules for mental health treatments, researchers do sequential random assignment of treatments; see Murphy (2005) for a general discussion and pointers to the literature. This design allows the researchers to have full knowledge of the assignment rule that generates the data they use to develop a statistical treatment rule and also to have plenty of clean, exogenous variation in the sequences of treatments received. The goal in this literature consists of adaptive rules in which subsequent treatments condition not only on initial treatments but on the outcomes observed in response to

those treatments. Broadly conceived, a similar goal exists in the context of non-deterministic sentencing regimes, wherein later stages, such as parole, condition their recommendation both on the initial sentence and on behavior in prison.

Aside from the obvious political and institutional obstacles to implementing multistage random assignment of punishments, other factors also render this option undesirable in a criminal justice context. In particular, while the incidence and type of new cases of mental illness do not depend on the treatment regime, the incidence and types of new crimes do. The set of crimes and of criminals that would appear under a randomized regime would differ from these under both the current regime and under a hypothetical deterministic statistical treatment rule due to differential sorting based on risk preference (as well as any differences in the amount and type of general deterrence generated by each system). Nonetheless, it is useful to keep multistage random assignment in the background as a benchmark of sorts.

One more realistic strategy thinks about undoing the effects of the existing implicit sentencing regime via “selection on observables”. That is, suppose that judges assign sentences based on some function of a set of observable characteristics plus a random term unrelated to recidivism risk. The latter might reflect the judge’s mood on the day of sentencing or his or her feelings about the defense attorney or the prosecutor and so on. The source of the variation does not matter, what matters is that observationally equivalent criminals receive different sentences for reasons unrelated to their probability of recidivism. Of course, this strategy requires that the researcher observe all of the variables that judges use in making their sentencing decisions that also correlate with recidivism. Existing data sets do not contain all of these variables, as the evidence in Gottfredson (1999) makes clear. Gottfredson collected information on judges’ subjective evaluations of the risk associated with particular convicted criminals and demonstrates

clearly that judges make use of variables that they observe but that do not appear in most data sets, and that the variables they use predict future behavior. Moreover, Gottfredson (1999) also makes it clear that different judges emphasize different goals in their sentencing decision, meaning that there is not just one sentencing regime, but many. This latter finding helps to motivate the existence of the required variation in sentencing conditional on observables that does not correlate with future behavior. The serious empirical pursuit of a selection on observables strategy for undoing the existing sentencing regime would require a substantial amount of rather creative additional data collection beyond that undertaken for existing data sets. Such a data collection enterprise would have many other benefits as well; not surprisingly, we heartily recommend it.

An alternative to selection on observables seeks to rely on sources of exogenous variation resulting from policy shocks or institutional design. The prisoner release exploited by Kuziemko (2006) and discussed in Section II constitutes an example of the former. Her paper also provides an example of the latter, as discontinuities in the existing Georgia risk assessment system allow her to obtain a clean estimate of the causal effect of time served for prisoners at the discontinuity points. Black et al. (2003) exploit a similar institutional feature in their experimental evaluation of the WPRS system. See e.g. Berk and de Leeuw (1999), van der Klaauw (2002) and Hahn, Todd and van der Klaauw (2001) for more on the statistical use of discontinuities.⁴

Each of the bits of exogenous variation exploited in the studies just described helps with the general task of undoing the effects of the existing treatment regime so that the effects of alternative treatment regimes can be consistently estimated. Fully accomplishing that task would require the accumulation of knowledge in this way for all aspects of the system, all types of

⁴ Also within this category of strategies based on exogenous variation lie single stage experimental exercises, such as that underlying the Philadelphia Bail Experiment analyzed in Abrams and Rohlfs (2005).

crime and all types of criminals. Like the data collection efforts required to make selection on observables work, this represents no minor undertaking. Achieving it would require a lot of flexibility on the part of administrators within the criminal justice system in order to allow researchers to introduce useful exogenous variation via policy shocks, random assignment or the creation of discontinuities in treatment. Like the data collection required for “selection on observables”, these activities would have substantial benefits for many other reasons; as such, we encourage them. They also hold the promise of allowing incremental improvements to parts of the criminal justice system, as with Kuziemko’s (2006) suggestion to restore the power of parole boards in light of the findings from her cost-benefit analysis.

VI. Conclusions

In this note, we have argued that researchers interested in studying risk must first undo the existing sentencing regime in order to allow a meaningful analysis of the potential of risk assessment based on observable characteristics and the associated use of statistical treatment rules in sentencing. Although this argument has not been made before in criminology, there have been related discussions. For example, the problem of selection bias in sentencing focuses on the fact that failure to observe decisions made by actors in the system to eliminate people from the sample (by dismissing charges, for example), limits our ability to make inferences about the impact of key variables such as race on sentencing outcomes (Bushway et al., 2007). Manski and Nagin (1998) demonstrate convincingly that different reasonable assumptions about the treatment allotment in sentencing juvenile delinquents can lead to radically different answers about the impact of prison on recidivism. With one assumed treatment scheme the empirical evidence supports a positive relationship between prison and crime while under another equally

plausible scheme the evidence supports specific deterrence. The problem they consider – estimating the causal impact of prison on subsequent behavior – represents just one component of the larger body of research required to meaningfully evaluate the potential of statistical treatment rules for high risk offenders.

The largely unrecognized conceptual difficulties with the current risk assessment literature imply that we need to start at the beginning if we want to assess the potential benefits of a statistical sentencing policy. The well-documented program of the Virginia Criminal Sentencing Commission hints at the potential value of such an effort. Other literatures that have addressed similar problems, both in economics and elsewhere, can provide useful guidance, though those contexts differ in important ways from that faced by criminologists. In our view, the time is ripe to abandon old conceptual frameworks (Auerhahn 1999) and begin work on a new round of research which pays explicit attention to the current treatment regime. This explicit orientation to current practice will not only allow us to directly inform current policy, it will allow us to make valid inference about the predictability of high risk behavior.

References

Abrams, D. and Rohlfs, C. (2006). Optimal bail and the value of freedom: Evidence from the Philadelphia Bail Experiment. Unpublished manuscript, University of Chicago.

Andrews, D., Bonta, J., and Wormith, S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency* 52:7-27.

Auerhahn, K. (1999). Selective incapacitation and the problem of prediction. *Criminology* 37:703-734.

Berger, M., Black, D. and Smith, J. (2001). Evaluating profiling as a means of allocating government services. In Lechner, M. and Pfeiffer, P. (eds.), *Econometric Evaluation of Active Labour Market Policies*, Physica, Heidelberg, pp. 59-84.

Berk, R. A., Ladd, H., Graziano, H., & Baek, J.-H. (2003). Randomized experiment testing inmate classification systems. *Criminology and Public Policy*, 2(2), 215-242.

Black, D., Smith, J., Berger, M., and Noel, B. (2003). Is the threat of reemployment services more effective than the services themselves? Evidence from random assignment in the UI system. *American Economic Review* 93:1313-1327.

Black, D., Galdo, J., and Smith, J. (2007). Evaluating the Worker Profiling and Reemployment Services system using a regression discontinuity design. *American Economic Review Papers and Proceedings* 97: PAGES.

Bushway, S., Johnson, B. and Slocum, L.A.. (2007). Is the magic still there? The relevance of the Heckman two-step correction for selection bias in criminology. *Journal of Quantitative Criminology* 23(2):151-178.

Engel, R. S. (Forthcoming). A Critique of the Outcome Test in Racial Profiling Research. *Justice Quarterly*.

Gendreau, P., Little, T., and Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology* 34(4): 575-607.

Gottfredson, D. (1999) Effects of judge's sentencing decisions on criminal careers. National Institute of Justice Research in Brief. NCJ 178889.

Gottfredson, S. and Gottfredson, D. (1986). Accuracy of prediction models. In Blumstein, A., Cohen, J., Roth, J., and Visher, C. (eds.). *Criminal Careers and "Career Criminals"*. National Academy Press, Washington, DC, pp. 212-290.

Gottfredson, S. and Gottfredson, D. (1994). Behavioral prediction and the problem of incapacitation. *Criminology* 32: 441-474.

- Gottfredson, S. and Moriarty, L. (2006). Statistical risk assessment: Old problems and new applications. *Crime and Delinquency* 52: 178-200.
- Hahn, J., Todd, P., and van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69(1):201-209.
- Harcourt, B. (2006) *Against Prediction: Profiling, Policing and Punishing in an Actuarial Age*. University Of Chicago Press, Chicago.
- Kleiman, M., Ostrom, B. and Cheesman, F. (2007). Using risk assessment to inform sentencing decisions for nonviolent offenders in Virginia. *Crime and Delinquency* 53(1): 106-132.
- Knowles, J., Persico, N., and Todd, P. (2001). Racial bias in motor-vehicle searches: Theory and evidence. *Journal of Political Economy* 109: 203-229.
- Kuziemko, I. 2006. "Going off parole: How the elimination of discretionary prison release affects the social cost of crime." Unpublished manuscript, Harvard University.
- Manski, C. and Nagin, D. (1998). Bounding disagreements about treatment effects: A case study of sentencing and recidivism. *Sociological Methodology* 28: 99–137.
- Murphy, S. (2005). An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine* 24(10):1455–1481.
- Ostrom, B., Kleiman, M., Cheesman, F., Hansen, R., and Kauder, N. (2002). *Offender Risk Assessment in Virginia: A Three-Stage Evaluation*. National Center for State Courts.
- Steffensmeier, D. (1980). Assessing the impact of the women's movement on sex-based differences in the handling of adult criminal defendants. *Crime and Delinquency* 26: 344–57.
- Steffensmeier, D., Kramer, J., and Streifel, C. (1993). Gender and imprisonment decisions. *Criminology* 31: 411-446.
- Steffensmeier, D., Kramer, J., and Ulmer, J. (1995). Age differences in sentencing. *Justice Quarterly* 12: 583–602.
- Tonry, M. (1996) *Sentencing Matters*. Oxford University Press, New York.
- Van der Klaauw, W. (2002). "Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review* 43(4): 1249-1287.
- Visher, C. (1986). The Rand inmate survey: A reanalysis. In Blumstein, A., Cohen, J., Roth, J. and Visher, C. (eds.) *Criminal Careers and "Career Criminals" Volume 2*. National Academy Press, Washington, DC, pp. 161-211.