REACTION ESSAY

THE DETERRENT EFFECTS OF CALIFORNIA'S PROPOSITION 8: WEIGHING THE EVIDENCE

STEVEN RAPHAEL
University of California, Berkeley

Whether, and the extent to which, stiffer sanctions deter crime is an integral question in fashioning corrections policy. To the extent that the behavior of the criminally inclined is completely insensitive to variation in punishment severity, crime-control policy should focus on incapacitation. Sentences should be long enough to ensure that offenders are removed from society during the period of their lives when they are most criminally active, but not so long as to be incarcerating individuals at a high cost to society who no longer pose a threat. On the other hand, if potential offenders take into account the severity of punishment when deciding whether to commit a felony (presumably, moderating their behavior in the face of stiffer sanctions), harsher sentencing may reduce overall offending, in part obviating the need to incapacitate those on the margin between offending and not offending. In the presence of such deterrence effects, optimal sentence lengths may extend beyond incapacitating offenders during the criminally active portions of their lives, as stiffer sentences will preclude some potential offenders from entering the system in the first place.

In practice, creating an effective and efficient sentencing regime requires precise knowledge of the importance of the deterrent effects of sentence length. Such knowledge, however, is hard to come by as empirical research on deterrence has faced several fundamental challenges that have been difficult to overcome.

First, the cross-sectional as well as longitudinal policy variation in sentence regimes often used to study deterrence is likely to be endogenously related to crime rates, a factor that could bias statistical inference in either direction. For example, if states with more severe crime problems enhance sentences as a result, cross-sectional analysis of the relationship between crime and sentencing severity is likely to reveal a positive relationship between sentence severity and overall crime. Alternatively, longitudinal analysis of crime rates within states may erroneously conclude that sentence enhancements reduce crime if states introduce harsher sentences in response to transitory increases in crime rates.

Second, assuming for the moment that policy variation occurring within a state is exogenous, correctly measuring deterrence requires assessing what would have happened in the state in the absence of the policy change, i.e., what's the counterfactual. Clearly, variation in crime rates (either across or within geographic areas) is driven by several determinants, including demographics, the economy, and corrections policy. The challenge to empirical studies of deterrence is to find a comparison region or regions with underlying crime fundamentals that are both similar at baseline and in time path to these determinants for the region experiencing a policy change.

Finally, identifying the deterrent effects of sentence severity must contend with the fact that the sentence itself incapacitates a criminally active person, and thus, it is likely to reduce to crime regardless of deterrence. In the ideal world, one would like to analyze how the behavior of criminally active individuals that are not currently incarcerated responds to changes in sentencing severity. In practice, however, most researchers have settled for analysis of aggregate crime rates that are theoretically influenced by both deterrence and incapacitation effects.

In their analysis of the sentence enhancements embodied in California's Proposition 8, Kessler and Levitt (1999) identify a clever strategy for addressing the third of the three identification problems listed above. Specifically, as Proposition 8 enhanced sentences for the most serious felony offenses, conviction of a Proposition 8 offense would result in a prison sentence even in the absence of the enhancement. In the short term, the proposition should have had no effect on the overall prison population, and thus, any measurable decline in crime relative to an adequately selected comparison group may be attributed to deterrence.

The critique of this work contained in the current study by Webster et al. (2006) indicates that Kessler and Levitt are less successful in addressing the first two challenges. In particular, the authors demonstrate that the pre–post-Proposition 8 decreases in crime appear to be continuations of trends that began after a spike in crime rates occurring two years earlier and subsequent declines that pre-date the proposition's passage. The authors also convincingly argue that the quasi-experimental comparison groups chosen are unlikely to accurately chart the counterfactual path that California crime rates would have followed in the absence of the proposition.

My reading of these two studies is as follows. Kessler and Levitt have offered a novel strategy for disentangling deterrence from incapacitation effects in focusing on crime changes in the immediate aftermath of sentence enhancements for serious crimes. Moreover, they make the valid point in the original article that this distinction is much more than an academic splitting of hairs, in that whether deterrence matters is key to setting optimal corrections policy. Nonetheless, the reanalysis presented in Webster et al. raises several important questions about the comparison groups employed by Kessler and Levitt. The reanalysis also raises important questions regarding the potential endogeneity of the proposition and the likelihood that crime would have likely declined in California in the absence of the proposition. In total, these questions cast serious doubt on the large deterrence effects reported in Kessler and Levitt.

In this brief reaction essay, I discuss in greater detail the identification problems faced by Kessler and Levitt (and empirical criminology research more generally) and offer some suggestions for refining their identification strategy and building on their original insight.

FILLING IN THE MISSING DATA

The identification problem faced in evaluating the effects of Proposition 8 is fundamentally a problem of missing data. The State of California underwent a regime change with regard to sentencing in June 1982. We observe crime rates before and after the implementation of the change. To truly identify the effect of the change in sentencing regimes, we would need to observe what would have happened in California absent the policy change. Of course, this data is unavailable as it is counter to fact, and thus, it is missing. The challenge to the empirical researcher then is to find a stand in for these missing data (the counterfactual) and then to compare the actual time path for the state with this hypothetical time path.

The crime series presented in Kessler and Levitt and the reanalysis by Webster et al. (2006) suggest at least two alternative identification strategies. First, one might exploit the discontinuity in California's sentencing policy and use the pre-policy levels and trends in crime rates to forecast what we would have expected to see in the absence of the policy change during the time period immediately after the passage of Proposition 8. In this instance, the results of such an analysis will depend critically on the years included, as demonstrated in Figure 2 of Webster et al. When only odd number years are presented, there is a visible discontinuity in crime rates surrounding the passage of Proposition 8, with murder, rape, robbery, and burglary rates increasing in all pre-proposition 8 years and abruptly decreasing in the post-change years. Using this presentation of the data, the pre-policy change trends in the data suggest increasing crime rates and the observed declines stand in stark contrast.

Filling in the even number years, however, we see crime rates peaking several years prior to the passage of Proposition 8 for three of the four targeted crimes and substantial decreases in crime preceding the policy change. In a regression-discontinuity framework, these higher frequency charts suggest that crime should have decreased through 1982 (to the extent that post-1980 declines represented secular trends in California crime rates). However, with only a few data points preceding the change, it is difficult to estimate with even a modest degree of certainty by how much crime would have fallen relative to the realized declines.

Although Kessler and Levitt do make note of the discrete declines in the odd years series, their principal estimators are not based on the time-series discontinuity for California alone, but on a comparison of California trends with a number of quasi-experimental control groups.[1] Specifically, the main results in Kessler and Levitt rely on (1) a comparison of pre–post-changes in crime rates for offenses with sentences enhanced by Proposition 8 relative to offenses with sentences not impacted by the proposition (a difference-in-differences estimator), and (2) a comparison of the relative change in Proposition 8 offenses to non-Proposition 8 offenses in California with the similar relative change for the rest of the United States (the difference-in-difference-in-differences estimator). To summarize their results, they find a relatively large pre–post-policy change decline in California crime rates for crime covered by Proposition 8 relative to crimes that are not covered (a negative difference-in-difference estimate) and a larger relative decline (covered relative to noncovered crimes) for California relative to the rest of the United States.

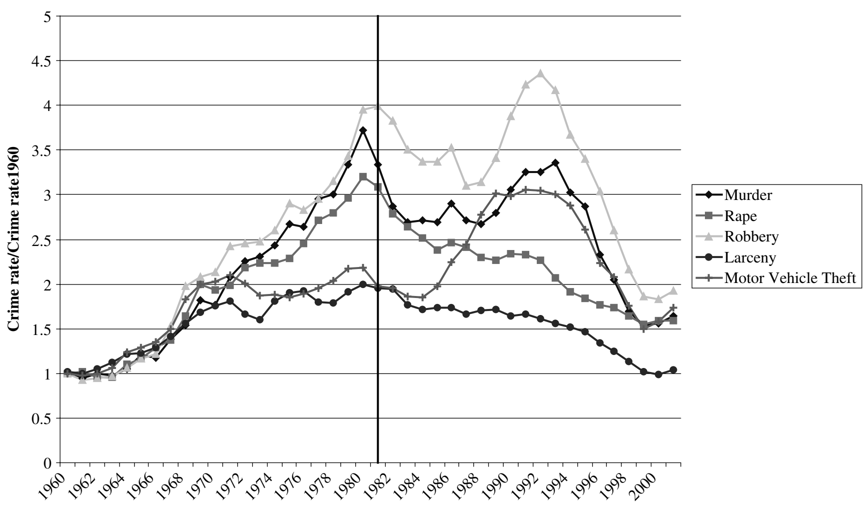
The validity of these results depends crucially on the quality of the quasi-experimental comparison groups chosen by the researchers. Two of the many questions one may ask of the comparison group are as follows: (1) Are the quasi-experimental treatment and control groups similar at baseline, and (2) have the comparison and treatment groups followed similar patterns in the past? The first criteria ensures that the researcher is making the proverbial apple-to-apple comparison, while satisfying the second criteria would provide some degree of assurance that the treated and nontreated crimes are influenced by similar factors. As Webster et al. note, by definition, Proposition 8 crimes are different at baseline (and at all other times) from non-Proposition 8 crimes both in terms of nature and frequency.[2] Thus, by the first criteria, the difference-in-differences strategy employed by Kessler and Levitt is problematic.

1. In the current context, this is an important consideration because many of the critiques raised by Webster et al. pertain to the time-series patterns in California crime rates in isolation. It may be the case that crime was already decreasing in California prior to the introduction of the proposition, but the policy change pushed crime rates below the counterfactual. For this reason, choosing an appropriate comparison group is key.

2. Crimes not covered by the proposition were considerably less serious felonies that occur with much greater frequency than covered felonies.

To assess whether the treatment and comparison crimes chosen by Kessler and Levitt tend to move together, Figure 1 presents crime rates for the State of California for the period 1960 to 2001 expressed relative to the crime rate in 1960 from the Uniform Crime Reports. Of the crimes presented in Figure 1, murder, rape, and robbery were covered in Proposition 8, whereas larceny and motor vehicle theft were not.[3] The figure reveals several important differences in the pre-Proposition 8 movements in these crime series. First, during the 1970s, there were much larger proportional increases in the violent felonies covered by the proposition relative to the nonviolent felonies that were unambiguously not covered. Moreover, the figure reveals greater variance in the covered crimes, with larger average relative increases and larger average relative decreases over the 40-year period analyzed. In conjunction with the critiques raised regarding the different natures of the offenses, these patterns suggest that the less serious felonies are a poor comparison group for the more serious crimes covered by the proposition.

FIGURE 1. CALIFORNIA CRIMES PER 100,000 RELATIVE TO THE 1960 CRIME RATE FOR FELONY OFFENSES



3. Regarding the other two part 1 felony offenses of assault and burglary, Proposition 8 enhanced sentences for certain assaults but not all, whereas only residential burglaries were covered under the proposition. For this reason, I omit these two crime series from Figure 1.

The second identification strategy employed by Kessler and Levitt (the relative change covered/noncovered in California relative to the United States) suffers from a similar lack of comparableness between the covered crimes in California and the exact crimes for the United States. The figures presented in the appendixes to Webster et al. reveal that for all crimes covered by the proposition, California crime rates pre-intervention were considerably higher than the comparable crime rates for the rest of the nation. In conjunction with the post-proposition declines in crime, the higher baseline crime rates for California relative to the United States is consistent with a transitory positive shock to violent crime in California that coincidentally died out with the passage of Proposition 8. Moreover, as the crime rates for the rest of the nation are averaged over a much larger population, one might argue that the volatility of crime cycles in California is likely to exceed the comparable volatility for the rest of the nation, as regional crime shocks are likely to be averaged out in the latter relative to the former. Similar to the within-state difference-in-differences comparison discussed above, the greater volatility of the California series suggests that relative year-to-year changes in California are always likely to exceed the similar changes for the rest of the nation, and they may spuriously create the impression of a larger relative decline in years corresponding to policy changes.[4]

In sum, the comparison groups chosen in Kessler and Levitt are problematic. The felonies not covered by the proposition are different in nature and follow different time paths in the 10- to 15-year period preceding the proposition. Moreover, for the serious felonies covered by the proposition, California crime rates far exceeded that of the averages for the rest of the nation, suggesting either that California was experiencing a transitory increase in violent crime or that the underlying fundamentals for the state were different.

Absent a satisfactory comparison group, I'm driven back to the basic time series for crime categories covered by Proposition 8. Inspection of the annual and monthly data presented in Webster et al. reveal that covered violent crime peaks and begins to decline sufficiently prior to the passage of Proposition 8. In the absence of further evidence, I must conclude that there is little evidence of a deterrent effect of this policy shock.

---

4. To be sure, Kessler and Levitt compare the covered–noncovered change in crime rates for California with the comparable relative change for the rest of the United States. Whether the greater volatility associated with averaging over a larger base for the United States exaggerates year-to-year variation in this difference-in-difference-in-differences estimator will depend on whether the differential variability between California crime rates and the rest of the United States is greater for covered crimes relative to noncovered crimes.

EXTENDING THE ORIGINAL ANALYSIS OF KESSLER AND LEVITT

The basic insight of Kessler and Levitt's identification strategy is quite novel and with some refinement may provide a solid estimate of the size and importance of deterrent effects. One extension of the Kessler/Levitt analysis may be to choose comparison states for California or for other states experiencing sentence enhancements through nonparametric matching on baseline crime rates and pre-intervention crime dynamics. Simple nearest-neighbor matches on pre-intervention outcomes can help identify states with the most similar pre-intervention experience to that experienced by the state undergoing the policy change.

An alternative identification strategy would be to move away from aggregate crime data and analyze the offending behavior of individuals subjected to alternative punishment severity. With individual-level data, one does not have to consider the impact of incapacitation as an institutionalization should be readily observable. Recent research by Lee and McCrary (2005) follows this path, using individual arrest data for youth with prior arrest records to test for a discontinuity in offending propensity upon reaching 18 years of age and coming under the purview of adult courts. The authors find little evidence of a discontinuous break in offending at 18 years of age.

Clearly, knowing the value of the deterrence parameter for the population of offenders likely to serve time would be of great value in setting policy. To date, the most compelling demonstrations of deterrent effects have relied on settings where nonincarceration sanctions have been made more severe for individuals who are likely to be quite different on average from the average felony offender in the United States. Thus, more research on the efficacy of deterring crime through variation in sentence severity is needed.

REFERENCES

Kessler, Daniel, and Steven D. Levitt
1999 Using sentence enhancements to distinguish between deterrence and incapacitation. Journal of Law and Economics 42:343–363.

Lee, David, and Justin McCrary
2005 Crime, punishment, and myopia. National Bureau of Economic Research Working Paper #11491.

Webster, Cheryl Marie, Anthony Doob, and Franklin Zimring
2006 Proposition 8 and crime rates in California: The case of the disappearing deterrent. Criminology & Public Policy. This issue.

Steven Raphael is Associate Dean and Associate Professor in the Goldman School of Public Policy at the University of California, Berkeley. His primary fields of concentration are labor and urban economics. Raphael has authored several research projects investigating the relationship between racial segregation in housing markets and the relative employment prospects of African-Americans. Raphael has also written theoretical and empirical papers on the economics of discrimination, the role of access to transportation in determining employment outcomes, the relationship between unemployment and crime, the role of peer influences on youth behavior, the effect of trade unions on wage structures, and homelessness.