THE CASE OF THE CRITICS WHO MISSED THE POINT: A REPLY TO WEBSTER ET AL.*

STEVEN D. LEVITT
University of Chicago and the American Bar Foundation

Webster et al.'s (2006) critique of my work with Daniel Kessler largely misses the point of the original article. In reading Webster et al., one is led to believe that our conclusions were based primarily, or even exclusively, on an analysis of the timing and magnitude of fluctuations in crimes covered by Proposition 8 in California before and after the law went into effect. This is simply not the case. We could not have been more explicit in our original article in emphasizing that inference based on that source of variation is likely to be highly misleading.[1] The results we report in the original article are based on comparisons of eligible and noneligible crimes, and the differential patterns of these sets of crimes inside and outside of California.

Webster et al. is an article that largely is devoted to criticizing us for a set of arguments we simply never made. Nothing in Webster et al. calls into question the analysis that we actually did carry out. When I replicate the original analysis, adding additional years of data or using monthly data rather than annual data, as Webster et al. suggest, the results remain completely consistent with the findings originally reported in Kessler and Levitt (1999).

WHAT KESSLER AND LEVITT (1999) ACTUALLY ARGUES

Kessler and Levitt (1999) began with the observation that the introduction of sentence enhancements provides an unusual opportunity to separately identify deterrence versus incapacitation. Because the criminal

* I would like to thank the editor, Phil Cook, for helpful comments. Marina Niessner provided outstanding research assistance. Author contact information: Steven Levitt, Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637 (e-mail: slevitt@uchicago.edu).

1. Indeed, it is the lessons learned from flawed past studies that relied solely on time-series variation to draw their conclusions (e.g., Zimring and Hawkings, 1993) that have been the impetus for recent empirical research to seek out more reliable sources of variation such as the "difference-in-differences" and "triple-difference" approach used in Kessler and Levitt (1999).

would have been sentenced to prison even without the sentence enhancements, there is no additional incapacitation effect from the enhancements in the short run. Therefore, any immediate decrease in eligible crime is likely to be attributable to deterrence. We then proceeded to analyze the impact of Proposition 8, passed into law by voter referendum in California in June 1982, which greatly increased the prison terms for repeat offenders committing "serious felonies."

The critical empirical challenge faced in identifying a causal impact of Proposition 8 on crime rates is that many other factors beyond Proposition 8 are, of course, simultaneously influencing crime. There are, for instance, clear national patterns in crime over this time period. Thus, it would be foolhardy to think that one could reasonably deduce a causal impact of Proposition 8 by simply comparing the difference in crime rates for eligible crimes before and after the passage of the law. Rather, our strategy involved a comparison of fluctuations in California crime rates for crimes covered by Proposition 8 and a set of similar (but less severe) crimes that were excluded from the law. This "difference-in-differences" approach only attributes to the law change the decline in eligible crimes in excess of the decline in noneligible crimes. Because there are many reasons why these less serious crimes might fluctuate in ways that systematically differ from the more serious offenses covered by Proposition 8, we took the further step of comparing the "difference-in-differences" estimates obtained in California with the same comparison for the rest of the United States. If the relative changes in serious and nonserious crime in the rest of the country mirror the changes that were taking place in California, then one would not want to causally attribute these changes to Proposition 8. Rather, in the "triple-difference" approach that we adopt as our preferred estimator, the impact of Proposition 8 is identified only off of any extra decline in eligible crimes in California relative to noneligible crimes in California, after taking into account how eligible and noneligible crime patterns were changing in the rest of the nation. The conclusion of our analysis was that Proposition 8 reduced eligible crimes by 4% in the year after passage and by 8% three years after passage.

Our original article could not have been more clear that any and all inference based on the time-series patterns in eligible crime is likely to be highly misleading. In the original piece we wrote, "Because Proposition 8 was passed by popular referendum, observed changes in crime around the time of its passage may reflect a combination of the true deterrent impact of harsher repeat-offender enhancements and of other factors correlated with but not caused by the law change, such as changes in demographics, in other state policies, and in broad social norms against crime. *This makes the availability of a control group of noneligible crimes critical to the analysis*." (Kessler and Levitt, 1999:355, emphasis added). In the next paragraph of the article, we note, "Inspection of Table 1 reveals that levels of crime rates in California are higher than those in the rest of the nation, but increases and decreases in California's crime rates tend to closely parallel those of the nation. . . . *Identifying a causal impact of Proposition 8 on eligible crimes in California requires differentiating between the impact of the law change and widespread decline in crime outside California that happens to coincide with its passage in 1982.* (p. 355, emphasis added). We then go on to write, "As the top row of Table 2 demonstrates, eligible crimes were rising in California before the passage of Proposition 8, then dropped sharply with the law change (a 17.5 percent decline) between 1981 and 1983, and remained roughly stable thereafter. *A naïve interpretation of the data might conclude that Proposition 8 had an enormous immediate effect that did not increase over time. Such a conclusion, however, is likely incorrect given the pattern of noneligible crimes in California (row 2)*. (pp. 355–357, emphasis added). Finally, we write, "Given that eligible and ineligible crimes exhibit systematic changes outside of California, the most convincing estimate of the true impact of Proposition 8 is the change in eligible crimes relative to ineligible crimes in California minus the corresponding change outside California. . ." (p. 357).

To emphasize a point that is already obvious from the preceding paragraph, none of the claims in our original article were based on the fact that eligible crimes in California only began to fall after the passage of Proposition 8. Because our analysis focused exclusively on comparisons between treatment and control groups, all that matters are the relative changes across these groups. If, as happens to be the case empirically, eligible crimes nationwide began falling a year or two in advance of Proposition 8's passage, that is not damaging to our story because our story was never based on the fact that eligible crimes fell in California. Rather, our arguments were based on the fact that after Proposition 8, eligible crimes fell *more* in California than noneligible crimes, and most importantly, the relative movements of eligible and noneligible crimes in California systematically differed from those in the rest of the United States after Proposition 8, but not before.

THE CRITICISMS OF WEBSTER ET AL.

The bulk of the criticisms in Webster et al. fall under the category of what they call "prima facie support," by which they mean an analysis of the simple time-series patterns in the data. As our conclusions were in no way based on the straight time-series patterns, this portion of Webster et al. is irrelevant in coming to a judgment about whether our original results are credible.

Webster et al. make three substantive arguments about the empirical work presented in Kessler and Levitt (1999). Their first concern is that we only report data for the odd-numbered years.[2] Their second concern relates to the use of annual as opposed to monthly data. Their third concern centers around the validity of the comparison groups and preexisting trends in the data. I deal with these three concerns in turn.

ADDING THE EVEN-NUMBERED YEARS TO THE ANALYSIS

The empirical estimates in Kessler and Levitt (1999) are contained in a single table (Table 2 of the original article). I have reproduced that table as Table 1 here, the only difference being that rather than using only the odd years of data as was done in the original article, all years of data are presented in Table 1. Each row of the table corresponds to a different crime category or to the difference across crime categories. The top three rows of the table report California data; the next three rows correspond to the rest of the United States. Crimes that are eligible for sentence enhancements under Proposition 8 (murder, rape, robbery, aggravated assault with a firearm, and burglary of a residence) are listed separately from noneligible crimes (aggravated assault with no firearm, burglary of a nonresidence, motor vehicle theft, and larceny). The final row is the difference between California and the rest of the United States. That bottom row is our estimate of the causal impact of Proposition 8. Columns in the table correspond to different sets of years that are analyzed, e.g., 1977–1981. The first four columns are various pre-Proposition 8 intervals. The remaining eight columns correspond to post-Proposition 8 periods.

The top row of Table 1 reinforces the fact that Webster et al. focus on—crimes eligible under Proposition 8 in California began falling in advance of the passage of the proposition. Between 1980 and 1981, eligible crimes fell by an average of 5.6% in California. A key point, however, is that noneligible crimes fell by a similar amount over that time period (6.1%). In the rest of the country, both eligible and noneligible crimes also fell between 1980 and 1981. Thus, our "triple-difference" estimate in the bottom row of the table is a 1 percentage point decline in California eligible

2. Although Webster et al. appear to cast our decision to present calculations based on the odd years of data in more sinister terms, the simple fact is that in treatment-control analyses of the sort we were presenting, it is standard to focus on longer intervals of time. Given that we were using annual data and the law change occurred in June 1982, the year 1982 was not readily classifiable as being either a "treatment" year or a "control" year. Consequently, the natural way to analyze the data was to look at preexisting trends over the period 1977–1981 or 1979–1981, and what happened after 1982, e.g., 1981–1983 and 1981–1985. Because of the "triple-differences" approach we adopt, there would be no variation left in any of the covariates that would typically be included in a regression analysis, and thus, there is no role for covariates in the analysis.

TABLE 1. ESTIMATES OF THE IMPACT OF PROPOSITION 8 ON CALIFORNIA CRIME RATES

	Pre-Proposition 8				Post-Proposition 8							
Geographic Region and Crime Category	1977-1981	1978-1981	1979-1981	1980-1981	1981-1982	1981-1983	1981-1984	1981-1985	1981-1986	1981-1987	1981-1988	1981-1989
California:												
Crimes eligible for Proposition 8	20.37	13.65	7.58	-5.57	-9.52	-17.46	-19.18	-20.59	-15.93	-19.83	-20.36	-15.49
Crimes not eligible for Proposition 8	9.48	3.67	-1.01	-6.11	-2.31	-8.60	-9.68	-7.23	6.43	9.12	13.71	17.77
California eligible - California ineligible	10.89	9.97	8.59	0.54	-7.21	-8.87	-9.50	-13.36	-22.36	-28.95	-34.07	-33.25
Rest of United States												
Crimes eligible for Proposition 8	21.09	18.87	7.99	-1.20	-9.28	-13.06	-14.68	-9.92	-4.12	-3.97	-5.06	0.27
Crimes not eligible for Proposition 8	11.16	8.53	-0.33	-2.74	-4.84	-8.03	-10.48	-4.15	0.81	4.40	4.77	12.26
Rest of U.S. eligible - Rest of U.S. ineligible	9.93	10.34	8.32	1.54	-4.44	-5.02	-4.20	-5.77	-4.93	-8.37	-9.83	-11.99
(California eligible - California ineligible) - (Rest of U.S. eligible - Rest of U.S. ineligible)	0.96	-0.36	0.27	-1.00	-2.77	-3.85	-5.30	-7.60	-17.43	-20.58	-24.24	-21.26

NOTE—Table entries are average percent changes in crime rates per 100,000 residents over the relevant crime categories in the years listed. Crimes eligible for sentence enhancements in California under Proposition 8 are murder, rape, robbery, aggravated assault with a firearm, and burglary of a residence. Ineligible crimes included in the table are aggravated assault with no firearm, burglary of a nonresidence, motor vehicle theft, and larceny. Values in the third row are the difference between rows 1 and 2. Values in the sixth row are the difference between rows 4 and 5. Values in the bottom row are the difference between rows 3 and 6. Proposition 8 took effect in June 1982. This table mirrors Table 2 of Kessler and Levitt (1999), except that even years have been added to the sample.

TABLE 2.
"TRIPLE-DIFFERENCE" ESTIMATES OF THE IMPACT OF PROPOSITION 8 USING MONTHLY CRIME DATA

Dependent Variable: Ln(Crime Rate)	
Impact of Proposition 8 (ß)	
1-6 months prior to Proposition 8	0.045 (0.053)
1-6 months after Proposition 8	-0.026 (0.060)
7+ months after Proposition 8	-0.100 (0.056)
California x eligible	0.376 (0.038)
California x post Proposition 8	0.017 (0.067)
Eligible x post Proposition 8	-0.099 (0.022)
California dummy	0.421 (0.038)
Eligible dummy	-0.273 (0.045)
Post Proposition 8 dummy	0.146 (0.048)
Forcible rape	1.025 (0.029)
Robbery	3.448 (0.024)
Aggravated assault with firearm	2.020 (0.030)
Aggravated assault with no firearm	4.055 (0.043)
Motor vehicle theft	3.963 (0.051)
Larceny	5.460 (0.045)
R^2	0.997
Observations	2,184

NOTE. -The dependent variable is natural logarithm of crime rates per 100,000 residents. See Eq. 1 and accompanying discussion for a description of the estimation approach. Crimes eligible for sentence enhancements in California under Proposition 8 included in the table are murder, rape, robbery, and aggravated assault with a firearm. Ineligible crimes included in the table are aggravated assault with no firearm, motor vehicle theft and larceny. This table includes monthly crime data on cities both inside of California and in the rest of the U.S from January 1977 until December 1989. Only cities with population greater than 200,000 in 1989 are included. Only actual crime reports are included. The omitted crime category in the table is murder. The estimated impact of Proposition 8 in the top 3 rows is relative to the period 7 or more months prior to the law's passage. Robust standard errors are in parenthesis. Standard errors are clustered by year x city. Proposition 8 took effect in June 1982. June 1982 is included in the "1- 6 months after Proposition 8" category.

crimes relative to the controls. If one looks over the two years in advance of the law change, 1979 to 1981, one finds a 0.27% increase in eligible crimes in California relative to the controls. The estimates for earlier years going back to 1977 are all close to zero. In other words, prior to the passage of Proposition 8, there were no systematic differences in the patterns of eligible and noneligible crimes inside versus outside of California.[3] It is this lack of a preexisting trend that gives credence to the methodology. Had there been large divergences before Proposition 8, it would be difficult to justify attributing variation after Proposition 8 to the law change when other factors must also be at work. At the risk of being redundant, the fact that eligible crimes began to fall before Proposition 8 is irrelevant to the analysis and conclusions in our article, contrary to the rhetoric of Webster et al. What matters is that the "triple-difference" estimates are close to zero before the law change.

The remaining eight columns of Table 1 demonstrate our estimates of the impact of Proposition 8 and how those estimates change over time. The top row shows that eligible crimes in California fell 9.5% in 1982 compared with 1981, whereas noneligible crimes fell only 2.3% in California. But eligible crimes in the rest of the United States also fell more than noneligible crimes. Thus, the estimate of the impact of Proposition 8 in the first six months after its passage is –2.8%.[4] Eligible crimes continue to fall faster than noneligible crimes over the ensuing years, both in California and in the rest of the country. But as the bottom row of the table makes clear, the relative declines are greater in California than elsewhere, and the estimated impact of Proposition 8 grows over time. This result is consistent with a role for both deterrence and incapacitation in reducing crime.

In summary, when one adds the even-numbered years to our analysis, nothing of substance changes. There continues to be little evidence of a preexisting trend, and there is a substantial and growing estimated impact of Proposition 8's passage.

USING MONTHLY DATA RATHER THAN ANNUAL DATA

Webster et al.'s second criticism of our analysis is the reliance on annual

3. In light of the lack of a preexisting trend, there is no concern that changes after the passage of the law reflect mean reversion. The exercise that Webster et al. carry out regarding mean reversion, because it is based off the raw time series data, rather than the "triple-difference," is irrelevant to the analysis in Kessler and Levitt (1999).

4. Because we are looking at annual data, we cannot tell whether this decline in crime occurred in the first half of the year, before Proposition 8 took effect, or in the second half of the year. It is for precisely this reason that in our original paper we omitted the year 1982 from the analysis and focused on changes in crime over the period 1981–1983, 1981–1985, etc.

data. As Webster et al. note, monthly offense data are available for only a select set of cities. It is also available for only a subset of the crime categories that we examine: murder, rape, robbery, assault with a firearm, other assaults, motor vehicle theft, and larceny. The reliability of these data is also more questionable than for the annual data (Maltz, 1999). Despite these concerns, I have compiled monthly reported offense data for the period 1977–1989 for the set of U.S. cities with a population greater than 250,000 as of 1989. Note that to carry out an analysis comparable with Kessler and Levitt (1999), one cannot focus solely on large California cities as Webster et al. did; one needs large cities in the remainder of the United States as well.

An important consideration when using monthly data is that they are extremely noisy. When I compute monthly estimates of the "triple-diffs" estimate, the standard deviation across these monthly estimates is 10.9%. Given that my estimate of the total impact of Proposition 8 in the first six months of passage is 2.8% (or less than 0.5% per month), it is clearly hopeless to look at monthly data and hope to discern a trend. The expected impact of 0.5% is less than one-twentieth of a standard deviation and is likely to be lost in the inherent noise in the process generating the monthly data. Because of the noise in the data, it will be similarly difficult to determine whether, in the months leading up to the passage of Proposition 8, eligible crimes in California were beginning to sharply deviate relative to noneligible crimes in a manner not seen elsewhere in the United States.

To address this question, I estimated regression equations that mirror the "triple-diffs" estimates reported in Table 1. I report regression coefficients rather than simply showing each of the "triple-diffs," both because with monthly data the potential number of "triple-diffs" estimates one might be interested in becomes extremely large and because of efficiency gains from imposing the restriction that the coefficients on particular sets of these "triple-diffs" estimates are identical. The basic form of a "triple-diffs" specification is as follows:

$$\ln(Crime_{itc})=\beta(Cal_i*elig_c*post_t) +\delta(Cal_i*elig_c)+\gamma(Cal_i*post_t)+\theta(elig_c*post_t) +\lambda Cal_i+\pi elig_c+\xi post_t+\varepsilon_{itc} \tag{1}$$

where i, t, and c index cities, time, and crime categories, respectively. The variable *Cal* is an indicator equal to one if the city is located in California and zero otherwise, *elig* is equal to one if a crime category is covered by Proposition 8 and is zero otherwise, and *post* denotes whether the time period falls before or after the passage of Proposition 8. The parameter of interest in Equation (1) is $\hat{a}$, which captures the differential response of eligible versus noneligible crimes in California relative to the rest of the country. The other variables in the regression capture other sources of variation that I do not want to use in identifying the effect. For instance, the coefficient λ absorbs any mean differences between crime in California and the rest of the United States over the whole period. Because the dependent variable is in logs, the coefficients are approximately interpreted as percent changes in crime.

The actual specification I estimate is slightly more complicated than Equation (1) because I do not want one estimate of $\hat{a}$ that reflects the average difference before and after Proposition 8, but rather, I want a particular focus on the path of crime in the months immediately before and after the passage of the law. The claim of Webster et al. is that my estimates are spurious because the changing crime patterns are predating the actual law change. To test this claim, I allow for separate $\hat{a}$ coefficients corresponding to the six months prior to the passage of the law, the six months immediately after the passage of the law, and the period more than six months after the law change. The omitted category is the period more than six months before the law change, so all estimates are relative to those earlier months of data.

The regression estimates are presented in Table 2. The key parameters are shown in the top three rows of the table. The top row reports the "triple-diffs" estimate of how eligible crimes in California relative to the control groups in the six months prior to the passage of Proposition 8. The coefficient of 0.045 implies that eligible crimes in California actually were 4.5% worse, relative to the period preceding that time. Because of the imprecision of the estimates, one cannot reject the null hypothesis of no difference between this six month period and the earlier times. Nonetheless, the positive coefficient on this variable is the *opposite* of what Webster et al. would predict. If anything, after controlling for the other factors, eligible crimes in California were rising relative to noneligible crimes and the rest of the United States. In contrast, in the six months after the passage of Proposition 8, eligible crimes fall by 2.6% relative to the earlier period. Again, because of large standard errors, one cannot reject the null hypothesis of no effect. Note, however, that this point estimate of –2.6% is almost identical to the magnitude of the –2.8% estimate reported in Table 1 for the first six months after passage using annual data. Similarly consistent with the annual data, I find that the mean impact of Proposition 8 for the period 7 months and beyond its passage is a 10% reduction and is highly statistically significant. Thus, although the monthly data are too noisy to precisely identify the impact of Proposition 8 in the months immediately before and after its passage, there is absolutely no evidence in these monthly data that the effects of the law were felt before it was passed and the point estimates match those obtained using annual data.

THE VALIDITY OF THE COMPARISON GROUPS AND PREEXISTING TRENDS IN THE DATA

The validity of any "difference-in-differences" approach rests critically on (1) the exogeneity of the timing of the law change in question, and (2) the availability of comparable control groups not affected by the law change. Webster et al. devote some space in their article (as we did in our original article) to asking important questions regarding whether the crimes excluded from Proposition 8 are an appropriate control for those covered by Proposition 8. The former crimes are less serious, raising important concerns. It is for that reason that the comparison of the patterns of these two sets of crimes inside and outside of California is so critical. Whether the rest of the United States is an adequate control for California is another question that can and should be debated. One fact that encourages us regarding the comparability of California and the United States is that, prior to the law change, the patterns in eligible and noneligible crimes inside and outside of California are very similar.

Although overall there are no differences in preexisting trends, as demonstrated in Table 1 here, Webster et al. focus on one particular pair of crimes (aggravated assault with and without a firearm) in which there is a difference in preexisting trends. Unfortunately for Webster et al., the particular example they chose makes our arguments stronger, not weaker. In the left-hand panel of Webster et al.'s Table 2, they show that using our triple-difference methodology, the estimated impact of Proposition 8 on aggravated assault with a firearm is –6.3%. In the right-hand panel of that same table, they also demonstrate that in the four years preceding the passage of Proposition 8, aggravated assault with a firearm increased more relative to aggravated assault without a firearm in California than in the rest of the United States. The typical concern when faced with a preexisting trend is that there are unobserved factors driving that trend and that those unobserved factors are likely to persist. The standard solution to this problem is to subtract off any preexisting trend from the differences observed after the law change, under the assumption that the preexisting trend would have continued unchanged absent the law change. In our context, this would be a quadruple-difference estimate. If one performs this quadruple-difference estimate, the impact of Proposition 8 on aggravated assault with a gun more than doubles to –13.7%. Not only does aggravated assault with a gun fall in California relative to the United States after the law change, but it also reverses a trend moving in the opposite direction. So if there is reason to think the trend would have persisted, the decline in aggravated assault after the law change is even more compelling.[5]

---

5. Despite this, Webster et al. argue that their Table 2 is evidence against our hypothesis rather than for our hypothesis. They argue that mean reversion should lead

CONCLUSION

Webster et al. (2006) show that crimes eligible under Proposition 8 began to fall before the passage of the referendum and continued to fall after. They interpret that finding as calling into question the conclusions of Kessler and Levitt (1999). No result of Kessler and Levitt (1999), however, depends on the downturn in eligible crimes in California occurring after the passage of Proposition 8. The entire analysis of Kessler and Levitt (1999) is predicated on the idea that many forces are influencing crime. Thus, to tease out any causal contribution of Proposition 8, one needs to identify *differences* in the pattern of eligible and noneligible crimes inside and outside of California. Neither adding additional years of data nor using monthly data, the two suggestions Webster et al. propose, change our results when one follows our methodology.

If we had claimed that a fall in crimes eligible under Proposition 8 was evidence of deterrence, and, if we had done this, missing the fact that such crimes were falling everywhere in the United States at the same time, that noneligible crimes were also falling, and that eligible crimes began falling before the passage of the law, then we would be fully deserving of excoriation far beyond what Webster et al. have put forth. But the simple fact is that we did not write such an article.

---

a crime that rose a lot in one period to fall in the next period. Although not inconceivable that such forces could be at work, there is no *a priori* reason to believe that the point at which mean reversion begins to operate is exactly the year of the law change (why not two years earlier or two years later, for instance) nor that mean reversion would be immediate and complete. We know that differences in crime across time and space can be very persistent; e.g., the U.S. homicide rate has been much higher than European homicide rates for many decades. There is little reason to think that crime processes are highly mean reverting.

If one truly did believe that the right model of crime was one of complete and immediate mean reversion, then the irony is that the raw time-series patterns observed in the data that are the primary criticism of our article by Webster et al. actually turns into *support* for a deterrent effect of Proposition 8. Crimes covered by the Proposition are falling in advance of the law change. By the Webster et al. mean reversion argument, in the next two years, these crimes would be expected to rise by as much as they fell in the preceding two years. The fact that crime continued to fall at the same rate would imply that the actual decline in crime was twice what was observed because absent the law change crime would have risen sharply. Thus, the logic used by these authors in the two different sections of the article is internally inconsistent.

REFERENCES

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

Maltz, Michael
1999 Bridging Gaps in Police Crime Data. Discussion Paper NCJ 176365. Washington, D.C.: Bureau of Justice Statistics.

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.

Zimring, Franklin, and Gordon Hawkins
1993 The Scale of Imprisonment. Chicago, Ill.: University of Chicago Press.