Rick Lempert

Growing Up in Law and Society: The Pulls of Policy and Methods

Richard Lempert

Eric Stein Distinguished University Professor of Law & Sociology, Emeritus, The University of Michigan Law School, Ann Arbor, Michigan 48109; email: rlempert@umich.edu

Annu. Rev. Law Soc. Sci. 2013. 9:1–32

First published online as a Review in Advance on August 14, 2013

The *Annual Review of Law and Social Science* is online at http://lawsocsci.annualreviews.org

This article's doi: 10.1146/annurev-lawsocsci-102612-134032

Keywords

social policy, Law and Society Association, social science methods, mismatch, affirmative action

Abstract

This article begins by tracing the aspirations and training that led to Lempert's commitment to the field of law and social science and includes comments on prominent figures in the field, the emergence of empirical legal studies, and other matters. It may interest scholars who seek to understand the history of the field's revival, and those who were among the first generation of Law and Society Association members may see some of their own experience in Lempert's account. The article then discusses policy uses of law and social science research and cautions against the possibility that a study's policy appeal may exceed the weight that can fairly be put on it. Five studies are used as examples: Wilson and Kelling's essay on "broken windows," Sherman and Berk's work on arrest for spouse abuse, Ehrlich's article on the deterrent effects of the death penalty, Lott and Mustard's work on right-to-carry laws, and Sander's mismatch critique of affirmative action. The article concludes by emphasizing the importance to policy of understanding mechanism and the need for sophistication in the soft methods of study design, along with a good understanding of formal statistics.

A PERSONAL HISTORY[1]

Dreams of the Law

I wanted to be a lawyer and attend Harvard Law School when I was five years old. I am not sure where these ambitions came from, but I was the proverbial Jewish boy who could not stand the sight of blood. I do know that my lawyer image was straight out of *Perry Mason*. I saw myself defending innocent men accused of murder and winning acquittals with my devastating cross-examinations and eloquence before the jury. Perhaps my mother's love of Erle Stanley Gardner communicated itself to me. As for Harvard, in my household as in the homes of many other first-generation Americans, it was the *ne plus ultra* in higher education, a view no doubt reinforced by my unlikely encounter with Emmett Harmon. Emmett, a Liberian who in 1950 was one of the rare black students to attend Harvard Law School, spent one of his spring breaks sleeping on our living room sofa. He and I, then seven-and-a-half, shared a fantasy that saw him becoming President of Liberia and me becoming President of the United States.[2]

A Turn Toward Sociology

I maintained my twin ambitions until my last semester at Oberlin College, by which time I had applied only to law schools and would soon be admitted by Harvard. My law practice ambitions fell victim to an Oberlin culture that admitted of only two respectable postgraduation paths—joining the Peace Corps or pursuing a PhD—and to the influence of Kiyoshi Ikeda, who taught me sociology. Kiyoshi had been a student of Harry Ball, one of the founders of the Law and Society Association (LSA). He had gotten from Harry and passed on to his students the belief that the sociology of law was the next hot area in sociology, especially for those who wanted to bring about social change.[3] Pivotal to my career was a reading course with Kiyoshi in my senior year, which drifted from discussing readings in the sociology of law to preparing a grant application for the investigation of an innovative Hawaiian program aimed at moving public housing tenants to homes of their own. Until then, I had never aspired to an academic career because, awed by the (perceived) brilliance of my professors, I didn't think I could measure up. Working with Kiyoshi, who treated me and others he mentored as if we were graduate students,[4] led me to believe that I might have what it took to be a successful academic. But by the time this realization hit, it was too late to apply to graduate programs in sociology, and having been admitted to the law schools at Harvard and Yale, I don't know if I would have gone on in sociology regardless. I chose Harvard over Yale largely because as a five-year-old, Harvard's prestige had led me to set my sights on going there, and after four years of living in Oberlin, I wanted nothing so

[1] I hope the reader will not be put off by the degree to which this article is autobiographical; an autobiographical focus is what I was told the editors wanted by way of introduction. Although my story is obviously personal, I expect that some who began academic careers in law and social science about the time I did will recognize aspects of their own careers in what I write. Others who are neither relatives nor close friends may wish to turn to page 11 where my focus changes from my life to observations on developments in law and social science over my career.

[2] Emmett came far closer than I to realizing this goal, as he became an important political actor in his home country. My child's understanding of how my encounter with Emmett came about is that, for business reasons, an uncle offered Emmett a place to stay during Harvard's spring break and then prevailed upon my parents, who at the time worked for him, to provide that place. Much later, I came to understand that although my uncle had a large house in Nutley, New Jersey, he did not think his neighbors would appreciate his giving even temporary hospitality to a black man. He did not see this as causing a problem in the all-white, working-class community where I lived.

[3] I was not the only one influenced by Kiyoshi's vision and enthusiasm. Over a period of five years, numbers of Oberlin graduates who regarded themselves as "Kiyoshi's students" went on to get sociology PhDs, with at least five of them, in addition to me, focusing on the sociology of law: Bliss Cartwright, Fred Dubow, David Ford, Robert Kidder, and Craig McEwen. In 2011, Kiyoshi received the LSA's Stan Wheeler Mentorship Award in part for teaching and inspiring this group. Several, like me, were not sociology majors and studied sociology only with Kiyoshi.

[4] I first met Craig McEwen when as an undergraduate Kiyoshi had him visit Michigan where I was a graduate student so he could to talk to Ed Laumann, a professor there, about smallest space analysis.

much as life in a city, and New Haven was insufficiently urban. My plan was to get a law degree and then seek support from the Russell Sage Foundation to study for a PhD in sociology, as I had heard that the Foundation supported social science PhDs seeking law degrees and lawyers pursuing PhDs.

A date during March of my year at Harvard changed this plan. My date, an Oberlin friend pursuing a PhD in government, told me she would be summering in Europe using money saved from her Woodrow Wilson fellowship.[5] I had fallen in love with France when I studied there the summer after my junior year at Oberlin, and I was immediately jealous. Moreover, although Erwin Griswold was paying my law school tuition,[6] I was dependent on my parents for living expenses, a situation I did not enjoy and they could ill afford. Thinking that I could have had a fellowship had I gone first into sociology, that I could be spending the summer in France, and that the Russell Sage Foundation would pay for my law degree if I had a PhD, I wrote to Kiyoshi the next day for advice. Following his suggestion, I wrote to Richard "Red" Schwartz at Northwestern, Leon Mayhew at Michigan, and Vernon Dibble at Columbia, telling each that I would be applying to his school to study the sociology of law and that I had to have full fellowship support to attend. I don't know what Kiyoshi wrote on my behalf, but although application deadlines had passed and I had never taken the GRE, within two weeks I had fellowship offers from Northwestern and Michigan.[7] Acting on my perception of the more prestigious alternative as I had when I chose Harvard Law School over Yale, I went to Michigan. To my retrospective regret, I never had the chance to be one of Red Schwartz's students.[8]

It turned out that my last days at Harvard had, in a way, been my best: I had aced my final exams and "made law review." This opened several doors. Harvard now appeared willing to accommodate my interest in studying sociology while I studied law,[9] and the Russell Sage

[5]The Woodrow Wilson National Fellowship Foundation paid what was then a standard amount, $1,800 plus tuition. A frugal graduate student could live on about $1,000 and travel abroad with what remained.

[6]Griswold, an Oberlin graduate and at the time both the dean of Harvard Law School and the chair of Oberlin's Board of Trustees, each year paid the tuition for one Oberlin graduate attending Harvard.

[7]I didn't hear from Columbia for months. It turned out that Dibble had been ill and had not received my letter.

[8]Mayhew at Michigan was, or would have been, an outstanding mentor. During my first year, Matt Silberman and I and several others were Mayhew's students in a seminar in the sociology of law. His book, *Law and Equal Opportunity* (Mayhew 1968), which we read, became for me a model of what good sociolegal research looked like. Mayhew soon left Michigan, however, and my plans to do a dissertation under his supervision never materialized.

[9]I had enjoyed my year at Harvard and the classes I attended, but there had been one jarring note. I was offered the chance to join Patricia Golden, a sociology graduate student, in a private reading course in the sociology of law with Talcott Parsons, then America's best known sociologist and the man whose theories motivated much of Kiyoshi's teaching. Knowing that Harvard's 2Ls and 3Ls could take up to two courses outside the law school for law school credit, I sought permission to take one of the two permitted courses in my first year. I was shocked when the associate dean told me that I could not take advantage of what seemed to me to be an incredible opportunity. His reason was that as a 1L, I had to work too hard. In fact, from the time I graduated high school until I finished my PhD, I never had an easier, less stressful year as a student than I did that first year at Harvard. Although the dean did not know it, I had during my first term audited a French course with substantial reading, including the first volume of Proust's *À la recherche du temps perdu*, and had no trouble keeping up with my law school classes. I responded to the dean's edict by taking the course with Parsons as an auditor, adding it to the French graduate seminar I was also auditing. Another response was to think Harvard Law School was hopelessly anti-intellectual. Today when many students do not see law review membership, even *Harvard Law Review* membership, as the apex of what a law student can aspire to, it may be difficult to appreciate how unusual my decision to depart Harvard after making law review was. Dean Griswold wrote me personally urging me to return, and I have somewhere a letter from the incoming president of the *Harvard Law Review* urging me to return to Harvard, explaining in great detail the opportunities I would lose by leaving and then concluding with a vague allusion suggesting that he wished he had had the courage to do what I did. My sister, a sophomore at Wellesley College, reported conversations on her campus referring to a student who had left Harvard after making law review, and years later, when in talking to John Langbein who began at Harvard the year after I did, I mentioned having spent a year there, his jaw dropped, and he looked at me and said, "You're the one!" I, perhaps naively, never felt that I was sacrificing anything nor that I was doing anything special. I still don't feel this way.

Foundation was willing, as an experiment, to support me while I studied simultaneously for the JD and PhD degrees. Academically I did even better at Michigan than I had at Harvard, perhaps because the competition was less stiff, and graduated with one of the highest grade point averages in the school's history, albeit only second best in my class. Before I even had my JD, I was invited to join the faculty, with my pursuit of a sociology PhD replacing the two or three years of large-law-firm experience that was the usual secondary credential that most young law professors then offered.

Starting at Michigan

At that time, I do not think there was a better place than Michigan to begin law teaching. Francis Allen, the nation's preeminent criminal procedure scholar and a marvelous person, had recently moved from the University of Chicago to become Michigan's dean. The school's center of gravity had moved from a more practice-oriented and conventional older generation with only a few members of genuine distinction to an outstanding group of younger scholars who would redefine the school in their more academic image.[10] There were always more slots available than we could fill, and as I was 25 when I was hired, for the next five or so years almost all our new hires were close to me in age and almost all became friends.

I had not known how good the school was when I accepted its offer,[11] but I soon learned that Michigan was not only one of the nation's top law schools; it was a school run by and for the faculty. Salaries were high,[12] I learned the joys of having a secretary, and there was substantial internal money to pay for research assistants and other research-related expenses.[13] The school even paid $1,200 to buy me a desk calculator that automatically calculated the standard deviation of a column of numbers. It was invaluable in the year before it became obsolete.

Perhaps best of all, tenure was virtually automatic. The decision was typically made in the third year of teaching and required only a

After I had decided to leave Harvard, another experience further soured me on the school. My roommate that year, a friend and debate partner from Oberlin at Harvard to study astrophysics, was getting married to another good friend in Chicago the day after my last exam, and I was to be the best man. I asked the assistant dean if I could take my last final a day early so that I could be in Chicago for the rehearsal dinner. This was not allowed. But the school was willing to mail the exam to an alumnus of theirs in Chicago and to allow me to take it beginning at the exact time it would be given at Harvard. The distrusting implication, that if I took the test early I might tell someone the questions, was obvious. I had gone to a school with an honor system that gave timed, take-home, closed-book exams, and I had never cheated nor did I know of anyone who had. I was greatly resentful. Only after I began teaching did I come to realize that some students do cheat, that it is impossible to tell which, and that I should have been grateful to Harvard for their willingness to accommodate my situation. It is a close question, but my experience still leaves me feeling that it is better to maintain an atmosphere of trust knowing some students will cheat than to see everyone as a possible cheater.

[10] Among the leaders of this transformative generation were Joe Sax, Terry Sandalow, Ted St. Antoine, Roger Cramton, Yale Kamisar, Arthur Miller, Jerry Israel, and an economist hired the same year I was, Peter Steiner.

[11] Terry Sandalow, who more than anyone else mentored me as a student and pushed the faculty to hire me, called me into his office after I had received my offer and cautioned me to think carefully about accepting it because I most likely could receive offers from other top law schools. I ignored his caution but have always been grateful that he gave it, and later I did the same with several of Michigan's top students who had Michigan offers.

[12] My starting salary level of $12,000 in 1968 was, I believe, more than my father had ever earned in a year. As my initial appointment was only 60% to facilitate my PhD studies, I received only $7,200 plus a summer stipend. The year before I joined the faculty, I took a teaching fellowship largely for the draft deferment that came with it. I also had my Russell Sage Foundation fellowship, which not only paid more than most fellowships but unlike almost every other fellowship at the time did not lower its support when a person had other sources of income. Add in summer earnings and the tax-preferred treatment of most of my sources of income, and what I took home as a first-year faculty member was less than my take-home pay the year before. It was not unusual for those entering teaching from law firms to take a salary hit, but I expect there have been few if any new law professors whose take-home incomes were less than what they had taken home as students.

[13] These resources had important career effects. They removed the pressure to apply for grants, so my research activities were usually small scale. Also they freed me to pursue whatever interests I had without worrying about funding, which is reflected in the diversity of topics I have written on.

single lengthy article.[14] Had I started my academic career in sociology, tenure would have required more writing of a higher quality.[15] I don't know whether I would have gotten tenure had a started in sociology rather than law, but I somehow doubt it, and whether I could achieve tenure as a young hire on Michigan's law faculty today is a matter I don't wish to contemplate. Rather, I am grateful that as a young faculty member, I felt almost no tenure pressure, with the reverse side of this coin being that I felt free to explore any issue that interested me and to work at my own pace.

I would be falsely modest if I suggested that hard work and intelligence were not in large measure responsible for my early (and later) career success. But it would be equally misleading to deny that luck and good fortune played a major role in how my career developed and perhaps that it even developed in the first place. I don't know where I would be if I had not taken Sociology 101 with Kiyoshi in my freshman year at Oberlin, or if a conversation on a date had not triggered a decision to leave Harvard Law School, or perhaps if I had not worked as a research assistant for Roger Cramton and if he had not been named chair of the law school's hiring committee for the year I graduated. But perhaps nothing illustrates my debt to good fortune more than my first publication.

I spent the summers of 1965, 1966, and 1967 in Honolulu, Hawaii, as a research assistant to Kiyoshi on the project I had assisted with at the proposal-writing stage. During my second summer, Donald Campbell, Northwestern's distinguished methodologist, visited Hawaii to teach a seminar on research design that Kiyoshi insisted that I and Bliss Cartwright, who was also working for him that summer, take.[16] The seminar, which for several of us continued for an hour or so in the cafeteria after its official end time, was mainly a seminar on Don Campbell, so it should be no surprise that my seminar paper was as well.[17] In it, I applied Campbell's designs for quasi-experimental research to legal impact studies. Campbell gave me an A, and I thought no more about it until a year later when back in Hawaii, I received a letter from Red Schwartz. The gist of the letter was, "I am editing a new journal, the *Law & Society Review* (*LSR*). Don Campbell showed me your paper on methods for legal impact studies. Do you mind if we publish it in our first issue?" Needless to say, I didn't mind. Thus I had my first publication—in the inaugural issue of the journal I was later to edit (Lempert 1966).[18]

Publishing the legal impact article had an unanticipated effect. It gave me a reputation in the law and society community and beyond as a methodologist, a reputation that has continued to this day but a mantle I have worn with considerable unease because of the implications of

[14]I received tenure midway through my fourth year rather than my third because there was an implicit understanding that tenure would await my completion of the PhD.

[15]Even before I could vote on tenure, I read my peers' tenure articles. I thought that several of these articles did not merit tenure even by the school's lax standards and that many tenure decisions were based less on the quality of the work submitted than on judgments about the quality of the candidate's mind. Given the selectivity of the hiring process, which allowed six or seven people with strong doubts to veto an appointment, it is no wonder that the quality-of-the-mind test was routinely passed. Perhaps more surprising is that for many years, the test worked. Although it is no wonder that the authors of strong tenure pieces succeeded, the authors of two of what I regarded as the three weakest tenure submissions I read were later hired by Harvard and Yale. The author of the third remained at Michigan and is widely regarded as a leading authority in his field.

[16]Hawaii's summer sociology faculty was at the time far more distinguished than its academic year staff because top scholars were happy to teach there in return for an expenses-paid summer in Hawaii. Ironically, it was there, rather than at Oberlin, that I met and became friends with Milton Yinger, Oberlin's most nationally prominent sociology professor.

[17]When I returned to Michigan after my summer of studying with Campbell, I sought to be excused from one of the Sociology Department's required methods courses because I felt that I had had the equivalent. When I approached Ed Swanson, who then directed the Department's graduate program, for his approval, he looked at me when I mentioned Campbell's name and said, "Don Campbell—the man has never written a bad article." I developed then and there the ambition that people would be able to say this about me.

[18]My next two publications, in the *American Sociological Review* (Lempert & Ikeda 1970) and the *LSR* (Lempert 1972a), were also course papers I gussied up for publication. At that point in my career, I thought that this publication business was easy—little did I know.

statistical expertise that go with it. Only I know how poorly grounded in statistics I am. I didn't take calculus until graduate school. Moreover, although my statistics teachers were not all bad, several were among the worst teachers I encountered, and I left their courses feeling I had learned little. I tried to compensate for being at sea in the mathematics of statistics by developing a good understanding of the assumptions and logic behind many statistical methods and the cautions that apply when using them. The little green volumes on statistical methods produced by Sage Publications were invaluable in this effort, and I expect that by the mid-1980s, at least 20 of them were sitting on my shelf. Unlike some of the other books there, the Sage volumes had been read.

A Year at Yale

Like many others of my sociolegal generation, my life, both academic and personal, was touched and made easier by the Russell Sage Foundation and by Stan Wheeler, who personified the Foundation's commitment to law and social science. The Foundation helped fund the research that brought me three summers in Hawaii and provided the connections that enabled me to do a dissertation on evictions from Hawaiian public housing. The thought that the Foundation would support my pursuit of a law degree if I had a PhD motivated my decision to leave Harvard Law School. The Foundation's generous fellowship support gave me a level of financial comfort attained by few of my fellow graduate students. And in 1971, Stan invited me to spend the 1971–1972 academic year at Yale Law School as a Russell Sage Foundation fellow.

This visit marked the beginning of my personal connections to the larger law and social science community. Getting to know Stan was a pleasure. Also, although I didn't see much of them during my visit, it was there that I first met Dave Trubek and Rick Abel, who were on the Yale Law School faculty. I also met Donald Black, whom I had not known at Michigan although he was only a year or two ahead of me and we shared a dissertation chair, Al Reiss. Don was in Yale's Sociology Department but would periodically visit the law and society wing of the law school building where, among other things, he spent some time in a futile effort trying to convert me to his view of the proper domain of the sociology of law (Black 1979). Bill Felstiner, then directing Yale's Law and Modernization program, also had an office in the law and society wing. He would in the second term audit a seminar I taught, and I recall reading and commenting on a draft of his important early article on social organization and dispute processing (Felstiner 1974). Apart from my Oberlin friend Bliss Cartwright with whom I had shared digs while we worked for Kiyoshi in Hawaii, the person I got to know best, because he and I were most similarly situated, was Bob Kagan who, even if seldom seen, remains a friend.

Being at Yale introduced me to another law school's culture. The contrast with Michigan was striking. In my first years at Michigan, I felt I learned more from coffee table conversation with my colleagues than I had in my law school classes. At Yale, faculty room conversation was less common and seldom involved more than two or three others. At Michigan, the dinner party was an institution.[19] By the time I had been on the faculty a year, I and my wife knew most if not all of the other professors and their wives, and we were friendly with many. At one of the few dinner parties I attended while at Yale, the wife of a faculty member of several years had to be introduced to other wives who were present. The most surprising and salient difference was, however, not social. It was my sense of how many Yale faculty, including some whom I had perceived as being among the law school world's biggest boosters of social science, regarded my discipline. Apart from Wheeler, Trubek, and Abel and with the

[19]I began at Michigan just before the women's movement took off. I think dinner parties flourished because few faculty wives—the faculty was then all male—had jobs outside the home. Dinner parties were a way of seeing others, and for some women, cooking was an outlet for creativity. As much as I enjoyed the faculty dinner party, I can't mourn its demise.

notable exception of Michael Reissman, the faculty most interested in social science seemed to value legal sociology as a tool that "real" law professors could use to inform their doctrinal and policy analysis but not as an enterprise to be valued in its own right. I wouldn't say that social scientists were disrespected, but they seemed to be regarded as scholars whose focus should be on matters that the faculty's "real lawyers" found of interest. Perhaps this attitude figured in Yale's decisions to deny tenure to Trubek and, more scandalously, to Abel.[20]

Yale's students were a different matter. While visiting, I offered a seminar titled "Selected Problems in Trials and Proof." Seldom have I so much enjoyed teaching. Five students plus an auditor (Bill Felstiner) took the course. Two of the students were children of Nixon Cabinet members, and two were on the upper staff of the *Yale Law Journal*. One of each, Nancy Rogers and David Kaye, went on to distinguished academic careers. Led by Kaye, we taught ourselves Bayes's theorem from the footnotes in Larry Tribe's (1971) article "Trial by Mathematics." Three of us, Daniel Kornstein (1976) in the *Journal of Legal Studies*, I in the *Michigan Law Review* (Lempert 1977), and Kaye in articles too numerous to mention, were early contributors to the post-Tribe literature on Bayes's theorem and the law. My article would not have been written had I not taught this seminar and, together with my students, wrestled with the question of whether, after Tribe's debunking of mathematics, Bayesian statistics had anything to offer the law.

Establishing a Reputation

A benefit of growing up in a newly developing field was that most of the giants of the field—I am thinking of people like Red Schwartz, Lawrence Friedman, Marc Galanter, Stewart Macaulay, Laura Nader, and Hans Zeisel—not only were still around but with the exception of Hans, weren't that much older than me. They have remained active scholars throughout my career; all are acquaintances, and most have become friends. Even better are the people one grows up with. So many people doing law and social science have enriched my life and have become dear friends that I could not begin to list them, but I also feel I must give a special shout-out to three: Shari Diamond, Felice Levine, and Joe Sanders; both my personal and professional lives are immensely better for having known them.

I finished my dissertation during my first term at Yale, received tenure at Michigan, and was recommended for promotion to full professor[21] shortly after my thesis defense. January 1972 was the beginning of the rest of my career. I didn't feel it began well. I had exhausted my publishable student-written papers and had only an insignificant paper or two I

[20]Trubek (1972) had published as his principal tenure piece a masterful explication of the core ideas in Max Weber's sociology of law. Until Anthony Kronman (1983) published his book on Weber, I think Trubek's article was the most accessible English language explication of what key portions of Weber's *Law in Economy and Society* were about, and even after Kronman's book appeared, I found that portions of Trubek's article were the clearest and best way to introduce sociology of law students to Weber's thinking. Nevertheless, I can understand how a law faculty might in good faith fail to appreciate Trubek's accomplishment and mistakenly regard it as more derivative than original. But no such excuse can be made for the decision regarding Abel, even allowing for the fact that in the early 1970s, Yale's tenure standards were the most rigorous in the elite law school world. Abel (1973) had written in article form a tome reviewing the literature on dispute institutions in society and attempting a theoretical synthesis. The work that had gone into this piece, the thinking behind it, and the originality of its contribution far exceeded the threshold that numerous Yale faculty had been judged to have exceeded when they had been granted tenure. Looking only at the quality of the scholarship, Yale's decision was inexplicable. It also had repercussions in other cases. I know of one Yale faculty member who missed getting tenure by one vote when at least one negative voter would have supported tenure had he not been outraged by the decision on Abel and used that decision as the standard to determine his later vote. I read that person's tenure article. It was not, by the standards Yale purported to apply, tenure worthy. The fact that its author came closer to getting tenure than Abel speaks both to the Yale faculty's failure to appreciate the social science enterprise and to the part played by "fitting in" in tenure decisions at Yale and elsewhere.

[21]There is nothing extraordinary here. Although the law school occasionally hired at the associate professor level, assistant professors granted tenure were promoted directly to full.

could pull from my dissertation. I did not feel myself drawn to any new topics, and although I did a neat, if little known, field experiment on pass-fail grading long before field experiments were common (Lempert 1972b), I began to doubt whether I could write anything of value that did not begin as a class assignment.

The closest I came to finding a topic that both attracted me and left me feeling I had something to say was when I read the Supreme Court's decision in *Williams v. Florida* (1970), which purported to rely on social science research for its conclusion that there was no discernible difference between the verdicts of 6- and 12-person juries. The studies it cited were, however, methodologically weak and even taken together could not support the Court's conclusion. So I contemplated as my next article writing a methodological critique of these studies. But I delayed too long. Before I could begin writing, an article by Hans Zeisel & Shari Diamond (1974) crossed my desk. Their article did what I had intended doing: It made clear the Court's mistake in relying on the *Williams* studies for the proposition that jury size did not affect jury verdicts. Zeisel and Diamond did not, however, stop there. They offered their own, better design for determining whether the verdicts of 6- and 12-member juries are likely to differ. Their design, however, had a subtle flaw. It failed to recognize that, as Kalven & Zeisel (1966) had shown a few years earlier, most of the cases that juries heard were not close, and there was no reason to expect that their proposed study would show juries of different sizes reaching different verdicts even if jury size could in close cases affect a jury's verdict substantially. Their design lacked the power to spot real-world differences.

Having been once preempted, I envisaged quickly writing a short article making just this point, but I became intrigued by the substantive issues of how likely it was that juries of different sizes would reach different verdicts and what those differences might be. Following a practice I was able to hew to for maybe 15 years until my time became too short and the extant literature too overwhelming, I read everything I could find that had any bearing on the question. For this piece, this meant reading the corpus of social psychological studies that contrasted group and individual decision making or decisions by larger and smaller groups. Not for the last time, I found that for the topic I had chosen, there was precious little directly on point. Studies comparing individual and group decision making were common, but the groups typically had no more than 3 members. Studies contrasting groups of different sizes were less common, and the larger groups in these studies seldom had more than 5 or 6 members. The studies I most wanted to be able to examine, those that compared decisions by 6- and 12-member groups, were as scarce as hen's teeth. Nonetheless, I did not reach a dead end, for the evidence as well as the implications of sampling statistics pointed in a consistent direction: Larger juries were likely to reach better decisions than smaller ones. The result of this effort was my article "Uncovering 'Non-discernible' Differences: Empirical Research and the Jury Size Cases" (Lempert 1975). As I did for an article I wrote on the death penalty (Lempert 1981), which together with "Non-discernible Differences" led several of the country's most prestigious law schools to communicate a strong interest in hiring me, I submitted this piece to both the *Harvard Law Review* and the *Yale Law Journal* and was turned down by each.

In some ways even more important to my career than my early social science articles was the book I wrote with Steven Saltzburg, *A Modern Approach to Evidence* (Lempert & Saltzburg 1977). This book pioneered the problem method of teaching evidence, for it broke with the then dominant casebook tradition by explaining doctrine textually and incorporating problems to test understanding.[22] What this book did for me, apart from later paying for

[22] At the time we wrote this book, every evidence book designed for classroom use taught the rules of evidence using extracts from cases. There had, however, recently appeared a problem pamphlet designed to be used in conjunction with the student edition of *McCormick's Handbook of the Law on*

my daughter's college education, was establish my reputation as a LAW professor. After publishing my evidence book, I did not have to prove to anyone that I could master legal analysis, write on legal issues, or teach law school courses. In this way, it freed me up to be a social scientist in a law school and to write only on social science topics. Although as it turned out I did write on court cases and legal doctrine from time to time, after my book appeared I never felt that my reputation as a legal analyst and evidence scholar depended on my doing so or that I had to do more to prove myself as a lawyer to my colleagues. In this I differed from a number of prominent law and society scholars who frequently complained that they felt somewhat marginalized on their faculties. Generalizing from my own experience, I would advise lawyer–social scientists who enter law teaching as junior faculty to write early on a significant doctrinal or analytic piece, and having established a reputation as someone who can play the legal analytic game, then forget about doing such work unless the spirit moves.[23]

The Urge to Matter

I wanted to become a lawyer because I wanted to make a difference in people's lives. As a five-year-old this meant saving innocent men from death. During my years at Oberlin, law's capacity to fight racism and promote economic equality were for me the profession's greatest attractions. During my first seven years of teaching, perhaps my greatest frustration was that I didn't think that beyond the narrow confines of the academy my work in any way mattered, and it didn't matter much within them. Even when the Supreme Court cited my jury size article in *Ballew v. Georgia* (1978) as support for the proposition that state criminal juries had to have at least six members, I did not think my scholarship had had an impact. As I told Stuart Nagel, who cornered me at an LSA meeting to report excitedly that we both had been cited in *Ballew*, being cited by the Court is usually nothing to crow about. Cited scholarship has seldom been read by the Justice citing it. It is invoked post hoc and inserted by a Justice's clerk to justify conclusions that would have been the same regardless. If it had any influence, its effect would have been indirect, mediated by its role in shaping a party's brief.[24] Later, however, I learned of legislative debates in which my jury article was mentioned, and I began to think my scholarship might have a real world impact.[25]

It is not, however, my scholarship that has done the most to satisfy my craving to matter. Rather, it is the opportunities to more directly affect consequential decisions that I would not have had but for my academic position and, in most cases, my writing. They began with Larry Ross's invitation in 1977 to serve on the National Science Foundation (NSF)'s Law and Social Sciences Program's initial grant advisory panel. Similar service has included membership on several National Research Council (NRC) panels (see, e.g., Normand et al. 1994), editorship of the *LSR*, the presidency of the LSA, and chairmanship of Michigan's Department of Sociology. The reader of this list will note that as I grew up, I set my sights lower than changing the world, but not entirely. I regard nothing I have done as more important or more satisfying than the role I played in *Grutter v. Bollinger* (2003) and my writings on affirmative action. As chair of the University of Michigan

Evidence. I used this combination the year before our book appeared, and it worked quite well.

[23]This advice may be dated by now, for over the past two decades, empirical analyses and social science have moved into the mainstream of law school scholarship.

[24]I confirmed this with a colleague who had been a clerk at the Supreme Court and had helped draft one of jury size decisions.

[25]I also learned that my jury scholarship, together with work done by Michael Saks and Shari Diamond, helped motivate the 1991 amendments to Rule 48 of the Federal Rules of Civil Procedure. Although the rule allows juries to have as few as 6 members (or as many as 12), it also provides that all sitting jurors shall participate in the verdict deliberations, and the Advisory Committee's comment urges courts to seat more than 6 jurors.

Law School's admissions committee, I took the lead in drafting the affirmative action policy at issue in *Grutter* and in steering it through the faculty approval process. Consequently I was a lead witness in the trial phase of this case. Moreover, two colleagues and I had surveyed Michigan's alumni to see how the school's minority graduates fared after law school, both absolutely and in comparison with the school's white alumni. We found no evidence that the minority alumni had done less well than their white counterparts (Lempert et al. 2000). The law school's attorneys chose not to use our study's results in developing their case, but the student interveners in *Grutter* called me as their witness. I was thus the only person to testify both for the law school, which defended its policy by arguing it was *Bakke* compliant (see *Regents of the University of California v. Bakke* 1978) and for the interveners, who sought to justify the law school's policy on grounds related to racial equality.[26]

My involvement in policy and my ability to have an impact increased dramatically when a second marriage led me to leave Michigan for Washington, DC, where I worked first for the NSF as division director for the Social and Economic Sciences and then for the Department of Homeland Security (DHS) as chief scientist in the Human Factors/Behavioral Science Division of its Science and Technology Directorate.[27] Outside of the academy, the NSF post was as close to a dream job as I could have imagined. Working for DHS immersed me in such bureaucratic nightmares and inefficiencies that I resigned my position two years before I was pension eligible. I have no regrets, however, for at NSF, I was still an academic, and at DHS, there was much new that I learned.

Community and Association

Throughout my career, the LSA has been my primary professional association. I attended its first annual meeting, organized by Red Schwartz when he was the dean at Buffalo Law School, and for several decades, I did not miss a meeting. I won't review my many involvements with the Association except to point out that in 1989, I was defeated by Rick Abel in what I see in retrospect as the most consequential presidential election in the LSA's history. As President, Rick spearheaded the Association's first international meeting, held in Amsterdam. Despite fears that the cost of travel abroad would dampen attendance, it proved to be the most widely attended meeting in LSA history. This was not only because many Europeans attended but also because many American law professors, perhaps attracted by the prospect of a school-paid trip to Europe, came even though they had not previously been involved in the LSA. What followed after the meeting did not change the Association's core membership or intellectual commitments but was in lesser ways transformative. Liking what they saw, some of these newbies joined the LSA, and others attended annual meetings if not regularly then from time to time. Membership in the Association grew, international attendance at the annual meeting increased, and income from the annual meeting became an important source of funding for LSA activities. Increased attendance also changed the character of the

[26] I do not claim any responsibility beyond the coincidental for the fact that Michigan's affirmative action program passed constitutional muster in *Grutter*. Not only was the policy a committee product, but had I not been appointed chair of the law school's admissions committee charged with drafting a *Bakke*-compliant policy, another faculty member would have had this task. Even if the result would have been a policy that differed in some particulars from what I drafted, I am sure it too would have been upheld. To the extent that any person deserves special credit for *Grutter*, it is the law school's then dean, Lee Bollinger, who realized that our then existing affirmative admissions process was vulnerable to attack and appointed a committee charged with revising it. The decision in *Fisher v. Texas* (2013), which some thought might overturn *Grutter*, has as I write today been handed down. *Grutter* is cited in *Fisher* with apparent approval, although the Court's opinion makes clear that it is assuming *Grutter* is still good law for purposes of deciding *Fisher*, rather than reaffirming the decision in *Grutter*.

[27] If one dedicated articles, I would dedicate this to Lisa, the woman I married and left Michigan for. She has never ceased to be a reason for smiles and gratitude.

annual meeting, with an increasing number of sessions organized around area or subject matter themes. One result has been more extensive interaction within areas of interest and less mixing across them. Another was that many of those newly attracted to the LSA had a lesser commitment to understanding law through empirical social science than those who founded or were originally attracted to the Association. Their involvement both broadened and diluted the LSA's intellectual core. This perhaps stimulated the growth of a separate empirical legal studies movement and led some Association members to treat the Society for Empirical Legal Studies rather than LSA as reflecting their primary academic identity. I have views about the desirability of each of these trends, but my point is not to evaluate them. Rather, it is to note that I would not have had Rick Abel's sense of the timeliness and importance of expanding the LSA's international footprint. Had Rick not been elected President, an international initiative might have been delayed for years, if it occurred at all. Clearly the time was ripe for international outreach, so I see the Association's voters as having made the right choice.

I expect my story could be told with some variation by many of those whose law and social science careers began in the late 1960s or early 1970s. Many were probably attracted to law at a young age and later became committed to social science. Many, no doubt, had a teacher or two without whom their lives would be very different. I know there have been literally dozens who were touched by Stan Wheeler and the Russell Sage Foundation. Those of my generation tend to know each other well. When we began, our numbers were small. We have for decades attended the same conferences, served on committees together, peer reviewed or edited each other's work, coauthored articles and books, occasionally disputed with each other in print, contributed to policy and developments both in our field and on the larger national stage, and formed friendships that are among the relationships we most value. And as my cohort has aged, I have been struck by how conventionally successful the people in it have been. Members of what was once regarded as a fringe field have become distinguished professors, department chairs, center directors, and deans, and prestigious schools have wooed people to new positions, even at ages when for most academics the likelihood of receiving an outside offer has dropped substantially.

THE GROWTH OF A FIELD

Research Capacity

Because members of my cohort began their careers almost with the birth of modern law and social science, we have a unique perspective on how the field developed. What follows is some of what I have observed, with a focus on research as it affects policy. My starting point is that when it comes to influencing policy, our sciences do matter, and they should. At the start of my career, social scientists, and not just those in law and social science, complained regularly that their findings were ignored in the halls of policy. We had a point. The evidence we gathered was often the best evidence available. But in retrospect, perhaps we complained too much, for the best social science evidence was often not all that good.[28] Important data collection efforts such as the General Social Survey (GSS) and the Panel Study of Income Dynamics (PSID) had not yet been launched or were in their infancy, as were ways of exploiting their repeated or longitudinal designs. Other data sources were limited, flawed, or one-time cross-sections. Statistical hazards, such as selection bias, were not just ignored, they were almost never recognized. Many of today's commonly used statistical methods had yet to be invented or if invented were

[28]It would nonetheless be interesting to systematically investigate the policy-relevant findings of the 1960s and 1970s and see which have held up and which have not. I would not be surprised if today's more sophisticated research confirms rather than counters much of what earlier researchers found—a tribute to the robustness of reality.

seldom employed. Advances in methodological understanding, statistical theory, and computer software and technology make the situation today quite different. Longitudinal data sets such as the National Longitudinal Survey of Youth (NLS), the Health and Retirement Study (HRS), and the PSID allow researchers to trace developments over time, within and across individuals, and in some instances, like the PSID, across generations, enabling confident distinctions between correlation and causation. Teasing out causation has also been facilitated by the use of true experiments to illuminate such law-related behaviors as policing (Braga et al. 1999, Boruch et al. 2000) and discrimination (Ayres 1991, Pager et al. 2009). Statistical techniques enabled by modern computing allow us to better model nonlinear relationships, including those that involve dichotomous and categorical variables. We have ways of coping with the problems posed by selection bias, and techniques of multiple imputation mean that missing variables do not hamper statistical analysis to the extent they once did. Bootstrapping, Monte Carlo methods, and improved approaches to meta-analysis allow relationships that once would have gone unobserved to be identified.

Technology and software have also transformed qualitative research. Observations may be visually recorded and archived. Transcripts and field notes can be more systematically recorded and linked. Not only are qualitative researchers better able to identify patterns in their data, but the curse of irreplicability has been partially lifted, for data once accessible only to the person who collected them can now be stored in ways that allow for easy review by other scholars. Given how difficult it can be to access some settings and the care needed to acquire and code qualitative information, these new capacities are of great importance. One result is a rapprochement between qualitatively and quantitatively oriented social scientists in fields such as political science and sociology, which were once riven by this difference (Brady & Collier 2010). These developments mean that although I haven't had a chance to read the essays that follow in this volume of the *Annual Review of Law and Social Science*, I can nonetheless confidently predict that among the studies the authors review, there will be some that use statistical methods and models that one didn't encounter a decade or two ago and others that systematically mix qualitative and quantitative methods with an adroitness in both approaches that few researchers until recently exhibited.

An increased appreciation for what social science has to offer has in turn enhanced our ability to access information that was once more difficult to acquire, if not completely off-limits. The change has been particularly important for law and social science because some of the more important settings where law is "enacted" are private, and second-hand reports of what occurs in these settings are no substitute for observation. Indeed, some venues that scholars have succeeded in entering seemed forever off-limits at the start of my career. Sarat & Felstiner (1995), for example, were able to sit in on meetings between divorce lawyers and their clients in 40 cases, following many from start to finish. Although much of what they observed was not inconsistent with the picture of divorce lawyering that social scientists since O'Gorman (1963) had painted, the view they provide is richer and more nuanced, describing lawyer-client interactions that other students of divorce had not seen. Even more surprising was Diamond et al.'s (2003) ability to record video of the discussions of juries deliberating actual cases. When Kalven and Zeisel attempted half a century earlier to record jury deliberations, the backlash was so intense that Congress passed a law making it illegal to record the deliberations of federal juries, and researchers assumed that states too would never permit this (US Congress 1955). Not only did the Arizona Supreme Court and the court in Phoenix approve of and in the case of the latter facilitate this project, but the idea for an observational study of jury deliberations originated not with the researchers but with the chief judge of the Phoenix court. Thus Diamond et al.'s effort[29] is pathbreaking in what it has done to illuminate the behavior of actual juries, and it also shows great respect for what students of law and social science have to offer the legal system, far greater respect than was prevalent throughout most of my career.

The Legal Academy

Changes in the legal academy have been as dramatic as those in data and methods. When I began teaching, few law schools had any empirical social scientists on their faculties, and the few empiricists who did teach in law schools often complained of being isolated or not respected.[30] Yale Law School, which was in the forefront of law schools interested in adding social scientists, hired two of the best, Red Schwartz and Jerome Skolnick, and tenured neither.[31] The only social scientists frequently found on law faculties were economists, but they were mainly theoretically rather than empirically oriented and typically relied on models laden with normatively questionable and empirically false assumptions, including the assumption that legal rules should aim at wealth maximization and that dollar-driven rational actor models accurately captured likely human behavior. Building on these assumptions led to a "law and economics" that most often produced normative recommendations that fell on the conservative side of the political spectrum, so much so that for many law professors, "law and economics" became synonymous with political conservatism in legal analysis. This too has changed. Treating wealth maximization as a normative goal is hotly contested, and the inadequacies of simplistic rational actor models are widely acknowledged. Moreover, the economists who inhabit law schools today are more likely to be empiricists than pure theorists, and like other law faculty, they cannot be pigeonholed politically, for their views range over the political spectrum.

The most profound change is not, however, the welcoming of empirical social scientists into the legal academy. Rather, it is the number of law professors, most without social science PhDs, who are doing empirical work. Their contributions have been facilitated by widely available data sets and software programs such as SPSS, SAS, and STATA, which not only enable sophisticated analyses but also are easy to use. The work being produced is not always of the highest quality. Some of it fails to understand the logic of social science inquiry or gives insufficient attention to model specification, data flaws, and the operationalization of crucial concepts. Moreover, research or theory that has not found its way into the law review literature is often ignored, and where it is considered, references are often to the most popular sources, blind to controversies within the disciplines (for a related discussion, see Lempert 2010). None of this is helped by the ease of publishing articles without the benefits of peer review in student-edited journals. But either by working with others or by educating themselves through programs such as the summer institutes in methods offered by Michigan's Inter-university Consortium for Political and Social Research (ICPSR), an increasing number of law professors without advanced social science degrees have nevertheless achieved a good understanding not only of statistical methods but also of how sound empirical inquiry proceeds. The result is that law reviews and law and social science journals regularly publish quality research done by "uncredentialed" law professors. Moreover, in doing their research, law professors, perhaps because they have so much bright student labor

[29]Other researchers, most notably Mary Rose and Beth Murphy, who were not part of the original data collection effort, were later integral to the project through their involvement in coding, data analysis, and writing.

[30]This was never my situation at Michigan, but I heard complaints from others, including some from distinguished scholars.

[31]Red Schwartz was later recruited from the Northwestern University sociology department to become dean of Buffalo Law School, apparently the first social scientist lacking a law degree to be appointed a law school dean. Jerry Skolnick later served on the faculties of the Boalt Hall (University of California at Berkeley) and NYU law schools. Either would have graced any law faculty.

they can call on, regularly assemble novel data sets that allow unique insights into issues (Beny 2008, Rehavi & Starr 2012).

If I have a regret regarding these developments, it is that different groups that should be constituents of a unified law and social science movement have in too large a measure gone their separate ways, and LSA, which should be the core professional organization for those interested in empirical inquiry into law, has never achieved its potential. This was to some degree true from the Association's start, when it took root among political scientists, sociologists, and empirically interested lawyers, with lesser representation from psychology and anthropology and almost no members from economics. The disinterest of economists in LSA is unfortunate but has, with a few notable exceptions, endured despite the Association's occasional, now mainly distant, efforts to court economists as members both by nominating some as trustees and by featuring economics at annual meetings. I see several reasons for this. Early on, the conservative implications of much law and economics research and the liberal tilt of the Law and Society Association may have played a role, but a more important and enduring reason is that enough economists were and are on law school faculties or otherwise interested in legal issues to support the publication of specialized law and economics journals and the formation of a law and economics professional association. Economists have not needed the Law and Society Association to help them find a professional home.

More unfortunate from my perspective is an opportunity LSA did not seize. The *Journal of Empirical Legal Studies* (*JELS*) should be an LSA publication, and law faculty doing empirical legal studies should see the Association as a comfortable home. Law professors and other scholars who regularly participate in the annual Conference on Empirical Legal Studies (CELS) but do not attend LSA's annual meeting would enrich the latter, and their presence within LSA would contribute to more fully ground the Association, its publications, and its meetings in the commitment to understand law through empirical research that it was once almost alone in championing.[32] Many people consider themselves members of both the law and society and the empirical legal studies communities; they present at both the LSA and CELS annual meetings, and they publish in both the *LSR* and *JELS*. But their numbers are too small. We do not need two separate law and social science associations, not when important synergies in executive office funding, publications, outreach, and training could be realized by a unified organization. Indeed, as journal-based modes of attracting and retaining members lose viability in the electronic age, it may turn out that consolidation will be the key to organizational survival.[33]

AFFECTING POLICY

At the start of the law and social science renaissance, Harry Kalven (1968) observed in an essay on the then young movement[34] that for social

[32]Let me be clear, I am talking about empirical research and not quantitative research. Empirical research uses both qualitative and quantitative methods, and both are important. *JELS* at its founding seemed to consider within its province only quantitative research, although more recently, receptivity to qualitative research has emerged. Nor should I be read as suggesting that the LSA is no longer committed to advocating and promoting the best empirical research on the law and to providing a home organization for social scientists and lawyers addressing empirically law-related issues. But my sense is that the Association's center of gravity, as evidenced by papers presented at the annual meetings, members elected to its Board of Trustees, and articles in the *Review*, has shifted so that the focus on rigorous empirical research, although still predominant, is less pronounced than it once was. Some will applaud this. I don't.

[33]Both the LSA and CELS have received substantial subsidies throughout their existence. The LSA has seen various educational institutions pick up a portion of the costs of running their Executive Office through various subventions, and CELS has enjoyed substantial law school support for its activities. If I recall correctly, when CELS held its first annual meeting, some of the legal world's best empirical researchers received monetary support to encourage their participation. To the extent subventions will be needed to maintain organizational integrity, over the long run CELS may be better situated than the LSA to survive because of the wealth of the law schools that employ its most prominent members.

[34]As a sign of its vitality, Kalven noted that he now had a "five foot shelf" of law and social science books in his office. I discarded the books from many such shelves when I moved

science research to influence policy, research results first had to become public knowledge. I believe the observation is for the most part accurate, but with an important qualification: The results of social science research may be influential even when the public within which findings are known is small and specialized. I recall no widespread public awareness of the research growing out of the Vera Foundation's Manhattan Bail Project, but the results became public knowledge in judicial circles and along with other studies helped persuade courts to dramatically change their bail-setting practices (Thomas 1976). Public knowledge of the research on hot spot policing is also low, but an increasing number of police departments are adapting their crime prevention strategies because of it (Braga & Weisburd 2010). Similarly, best ways to conduct lineups are not common knowledge, but the research of Gary Wells and others has led to a substantial revamping of criminal identification practices in some jurisdictions (e.g., Wells & Penrod 2011).

But even without my qualification, Kalven was on to something. Often it appears that social science research or theory must capture the popular imagination—usually in the form of a pithy take-away message—before one sees an influence on judicial decisions, legislation, or systemic reform.[35] I am sure I am not alone in feeling that I have done research with important policy implications that, never having reached audiences outside the academy, caused nary a ripple in any larger social context. When social science research does cause ripples, it can be for better or worse, depending on the quality of the science and how well it is understood by policy makers. But it is not always the quality of research that determines its reception. The dissemination of what is learned is crucial, and some researchers do far more than others to call attention to their findings. Moreover, ideological or other interest groups regularly tout research results they find congenial, regardless of a study's quality or the reliability of synthetic analyses.

Five Uneasy Pieces

"Broken windows." Perhaps it is only because it is the field I know best, but I think that law and social science has had more than its share of studies that resounded in the halls of policy when at best they raised issues for further examination and at worst seemed aimed at promoting particular policies regardless of weaknesses in the science. An example of the former is James Q.Wilson & George L. Kelling's (1982) thought-provoking "Broken Windows" article. This was never a work of empirical social science; rather, it was a suggestive hypothesis, and the authors were clear on this score. Still, perhaps because it appeared in a widely read magazine or perhaps because it appealed to commonsense notions of how people might behave, it was treated early on as if it provided a proven prescription for reducing crime. The evidence on whether reducing visible signs of neighborhood disorder prevents more serious crime is mixed, but the research most supportive of this approach often appears to confuse correlation with causation or reports results that are more likely traceable to an increased police presence or other interventions that go beyond fixing "broken windows." The most systematic study focusing on the kinds of visible signs of neighborhood disorder that are of a piece with broken windows fails to find in visible signs of disorder an important cause of serious crime (Sampson & Raudenbush 1999, 2004).[36]

from Michigan and still lacked shelf room for the volumes I brought with me.

[35]The requisite popular imagination need not be widespread. Because the legal system's decision makers are most often elites, knowledge that is largely confined to elite publics may nevertheless be greatly influential.

[36]Sampson & Raudenbush (1999) find that neighborhood disorder may have a small effect on robbery, and they leave open the possibility of indirect effects. They write:

> What we would claim, however, is that the current fascination in policy circles on cleaning up disorder through law enforcement techniques appears simplistic and largely misplaced, at least in terms of directly fighting crime. Eradicating disorder *may*

The spouse abuse experiment. The Minneapolis Domestic Violence experiment, which sought to assess the effects of arrest on recidivist spouse abuse (Sherman & Berk 1984), is another example. Unlike Wilson and Kelling's speculative suggestions, Sherman & Berk (1984) were reporting a study's results, those of a nicely designed field experiment that indicated that arrest for misdemeanor spouse abuse deterred future domestic violence. The study had its problems, including some compromising of the randomized treatment design, results that differed with the choice of dependent variable, and little attention to mechanism. Nevertheless, the study was well done and almost unique in its time. Its impact was substantial, as it was used to persuade police departments to abandon counseling approaches to domestic violence in favor of mandatory or presumptive arrest polices. The impact is in part traceable to the era, for Sherman and Berk's findings came at a time when feminist concerns to get tougher on domestic violence coincided with the push by law-and-order conservatives to get tougher on all crimes. It was also no accident that the study's results became public knowledge. Larry Sherman, the principal investigator, went to extraordinary lengths to disseminate his results, including filming police action during a ride along in order to provide television news shows with video to accompany reports of the study's results (Lempert 1989, Sherman & Cohn 1989). The publicity and policy influence were, however, premature. As Sherman himself discovered, the Minneapolis results did not consistently replicate in other cities. Rather, the effects of arrest on recidivist abuse seemed to depend on the characteristics of abusers and on how the legal system treated men after arrest (Sherman et al. 1992a,b).

The replications also reinforced a feature evident in the Minneapolis data. Conclusions depended on how recidivism was measured. Thus, a later analysis that combined the data from five replication sites found a statistically significant effect suggesting deterrence if repeat abuse was measured by later spousal interviews but found no statistically significant relationship if police records were the measure of repeat offending (Maxwell et al. 2001, 2002). Although the results from this global analysis suggested arrest might prevent abuse and was unlikely to increase it, from a policy standpoint, their implications are ambiguous. Despite the care taken by the researchers, questions of what to make of the core results remain. Although the spousal interview data suggest substantial effects, about a quarter of the spouses[37] were not reinterviewed, and those who were reinterviewed spoke to the researchers differing numbers of times and at different time lags with respect to the arrest. Thus, the interview sample could be biased in unknown ways. Moreover, the interview data lump together subsequent violence with what the authors call "aggression," or the unconsummated threat of violence. These concerns are made more serious because the police rearrest data, although in the deterrence direction, did not achieve statistical significance. In addition, 248 cases involving repeat suspects, that is, men who had previously been subject to an experimental treatment, are excluded from the analysis, but these may be the cases that involve the greatest threats of serious violence.

Even more instructive is the difficulty of deriving policy from the reanalysis assuming its core findings are not problematic and we can conclude that arrests in the aggregate were never associated with increased reoffending

indirectly reduce crime by stabilizing neighborhoods, but the direct link as formulated by proponents was not the predominant one in our study. What we found instead is that neighborhoods high in disorder do not have higher crime rates in general than neighborhoods low in disorder once collective efficacy and structural antecedents are held constant. (p. 638)

[37] The combined analysis looked only at female victims, although some of the studies included incidents in which women were the aggressors. The female victims were often not actual spouses because they weren't married to their abusers, but for convenience, "spouse" is the label applied to all female victims.

and in some cases reduced or postponed future aggression.[38] Problems arise because the effects of arrest seem modest; the treatment received was less important in explaining subsequent violence than were the implications of such other suspect characteristics as age and criminal record, and the majority of those assigned to treatments other than arrest did not reoffend, whereas some who were arrested were chronic reoffenders. Given the substantial costs arrest can impose on both the offender and his family, it is not clear that presumptive or mandatory arrest is, even from the perspective of a woman who might enjoy somewhat greater protection, a wise policy. The randomized arrest experiments offer little aid in answering this question. They were designed only to measure costs that took the form of future violence. Neither the Minneapolis study nor its replications sought to measure other kinds of costs or benefits that the treatments studied might impose.

Ehrlich and the death penalty. The preceding examples involve, in one case, an intriguing idea well worth testing and, in the other, well-designed innovative research that was prematurely influential. But not all influential studies have such strengths. Serious scientific shortcomings have too often failed to derail the policy influence of law and social science research. One of the most prominent studies that may be so characterized nevertheless made a genuine contribution. I am thinking of Isaac Ehrlich's (1975) longitudinal study of capital punishment. Ehrlich's valuable contribution was not, however, to the policy debate. Rather, he was the first to use modern time-series econometrics to investigate the deterrent effects of capital punishment, and for better or worse, he refocused the debate from the effects of death penalty laws to the effects of actually executing. His conclusion, that each execution saves eight lives, was, however, seriously flawed because as others soon showed, it depended entirely on his choice of years to analyze, and the last years included in the time series were unusual because for different reasons, executions dropped to almost zero and crime of all types, not just homicides, increased dramatically. If executions in fact deterred, then one should have seen a deterrent effect even if the last five or so years in the time series were dropped. But eliminating these years eliminated any suggestion of deterrence. Indeed in some reanalyses, the association between executions and homicide was positive (Lempert 1981).

More guns, less crime. Two other studies have made even less of a contribution; indeed, if they have stirred up the pond, it is only to muddy the waters and perhaps promote unwise policies. One, by John Lott & David Mustard (1997), purports to show an inverse relationship between the existence of right-to-carry laws and the presence of violent crime. But the study contained coding errors, and the results were heavily dependent on subjective judgments about model specification. When the coding errors were corrected and different specifications employed, results were quite different and often not consistent with the authors' claims (Ayres & Donahue 2003). The controversy surrounding this work eventually led the NRC's Committee on Law and Justice to devote a chapter of its book *Firearms and Violence: A Critical Review* (Wellford et al. 2004) to closely examining Lott and Mustard's work and other related studies in order to ascertain whether there was empirical support for right-to-carry laws sufficient to guide policy. After reviewing the existing studies and conducting its own data analysis, the Committee concluded

> that, in light of (a) the sensitivity of the empirical results to seemingly minor changes in model specification, (b) a lack of robustness of the results to the inclusion of more recent years of data (during which there are many more law changes than in the earlier period), and (c) the imprecision of some results, it is impossible to draw strong conclusions from

[38]The combined analysis indicates that there was a slightly longer time to reoffending among the arrested group than among those subject to treatments other than arrest.

the existing literature on the causal impact of these laws. (Wellford et al. 2004, p. 121)[39]

Suspicions about the accuracy of Lott and Mustard's results did not, however, have to await the results of NRC review or even the attempts to replicate the Lott and Mustard models and test their sensitivity to different specifications. The results of the original study were sufficiently puzzling that even a reader unfamiliar with the methods used should have been reluctant to give them much credence. The major puzzle was that in addition to indicating that right-to-carry laws were associated with statistically significant drops in certain violent crimes, the model showed statistically significant increases in such property crimes as larceny and auto theft (Lott & Mustard 1997, p. 20). Lott and Mustard treat this as an explicable if not expected substitution effect—namely, people deterred from killing or raping by the existence of right-to-carry laws decide to steal from stores or take autos instead. I am tempted to say that only an economist could believe such a theory, but I don't think that most economists, especially economists versed in criminology, would believe that such differently motivated crimes (Katz 1988) plausibly substitute for each other. Moreover, it is hard to understand why knowing someone might have a gun would deter someone from attempting to kill him but not from taking his car or stealing from his store. The model's weakness is also suggested by significant coefficients on variables that don't make sense. For example, controlling for the presence of black males of various ages, the proportion of black females between the ages of 40 and 49 has no relationship to a county's overall violent crime rate but is related significantly, and highly so, to the county's homicide, rape, aggravated assault, and auto theft rates (Lott & Mustard 1997, p. 21). When parts of a model's output make no theoretical sense, one should hesitate before putting too much faith in what the model says about key dependent variables, even if a theoretically plausible story can be offered to explain them.

It is no accident that the results of the above studies in some quarters became popular knowledge. Not only did many people find Lott and Mustard's results intuitively plausible, but the results supported policies that many people favored. Compatibility with people's intuitions or desires is not, however, the full explanation for why the results of these studies became widely known. Wilson & Kelling's (1982) theory was well known from the start because it began life as a story in a widely read popular magazine. Sherman, as I noted, took extraordinary steps to publicize the results of the Minneapolis experiment (Sherman & Berk 1984), including making the topic television friendly. Ehrlich (1975) made the news when he testified before Congress about his findings even before they were published, and his article was cited for the proposition that the deterrent effect of the death penalty was in dispute when the Supreme Court reinstituted the death penalty in *Gregg v. Georgia* (1976).[40] The Lott and Mustard article was, not surprisingly, music to the ears of the National Rifle Association, which ensured the wide dissemination of the authors' results.

Mismatch and affirmative action. The same is true of the last of the studies I shall discuss,[41]

[39]One member of the committee, the late James Q. Wilson, dissented to this conclusion with respect to homicides, but no other member agreed, and in my view, the committee effectively answered Wilson's dissent [see appendixes A and B in *Firearms and Violence* (Wellford et al. 2004)].

[40]Justice Stewart used the existence of this dispute to justify allowing Georgia to decide for itself whether the death penalty served a proper purpose without considering whether it could serve a proper purpose if the penalty did not deter. But the fact of a dispute was supported only by a citation to Ehrlich. The other studies the Justice cited either criticized Ehrlich or presented evidence suggesting that the death penalty did not deter (*Gregg* 1976, pp. 184–87).

[41]I could offer other examples of research that became popular knowledge and influenced policy despite the unreliability of its findings. Two that come to mind are Robert Martinson's (1974) influential summary in *The Public Interest* of a wide-ranging survey he and others had done of programs for in-prison rehabilitation. The pithy conclusion, "nothing works," which helped fuel the movement away from the

those conducted by UCLA law professor Richard Sander using data mainly from the Law School Admissions Council's Bar Passage Study (Wightman 1998) and offered by him and others to challenge the wisdom of using affirmative action to boost minority, and especially African American, law school enrollments. Although I have in various articles commented on several of the studies I mention above, none has engaged me so much as Professor Sander's work. His argument is that affirmative action does not help the minority students who are its supposed beneficiaries because it places them in situations of institutional mismatch. What follows, according to Sander, is that students admitted through affirmative action find themselves in schools where they cannot keep up with the level at which courses are taught; as a result, they learn less than they would have learned had they had attended less selective institutions. Their learning difficulties mean they are more apt to drop out and, in the case of law students, to fail the bar exam should they manage to graduate. So serious are these detriments that Professor Sander claimed that the likely result of law school affirmative action is to reduce the number of African Americans who succeed in becoming lawyers (Sander 2004, Sander & Taylor 2012).

As with the other work I mentioned, it is no accident that Professor Sander's mismatch thesis has received considerable public attention. His claims are intuitively plausible and have great appeal for those opposed to affirmative action. Not only are they purportedly based on solid scientific evidence, but as advanced, they offer little ground for minorities to cry racism. Sander forthrightly states that African Americans perform as well as whites or Asians with similar academic credentials, and he claims his work is aimed at helping minorities avoid serious mistakes in their choice of law schools and, by extension, other institutions of higher education. As three coauthors and I stated when critiquing his work, if we believed the findings he touts, we too would have questions about affirmative action (Chambers et al. 2005). But it is not Sander's arguments alone that have brought his work to public attention. Rather, its dissemination bears the mark of public relations professionals. *Mismatch* (Sander & Taylor 2012), the book summarizing Sander's arguments that he wrote with Stuart Taylor, was timed to appear coincident with Supreme Court oral argument in *Fisher v. Texas* (2013; see n.26 above), the latest case challenging the constitutionality of affirmative action. Shortly before its official publication date, the book was featured at a widely attended forum at the Brookings Institution with all the speakers, some of whom were Sander's coauthors, chosen by Sander and Taylor. No prominent critic of Sander's work was on the program. Other outreach efforts included op-eds in several papers including the *Washington Post* and *Wall Street Journal*, an interview on NPR, blog postings, and mentions by conservative columnists.

Sander's work is, however, more likely to mislead policy makers than to help them. Several coauthors and I looked closely at his data and methods. We found flaws in his statistical analyses and conclusions that did not stand up to close scrutiny (Chambers et al. 2005).[42] We were not the only social scientists to reach this conclusion. With the exception of several papers by a longtime Sander coauthor,[43]

rehabilitative ideal to a retributionist perspective, was later shown to be overstated, as Martinson himself came to recognize (Martinson 1979, Gendreau & Ross 1987). A second is Charles Murray's (1984) *Losing Ground*, which greatly influenced welfare policy in the United States but was regarded by leading experts as seriously flawed in its reading of the extant literature (Danziger & Gottschalk 1985, Jencks 1985).

[42] Professor Sander (2005) replied to our critique and that by Ayres and Brooks. For reasons I and my coauthors documented in detail, his response was unpersuasive (Lempert et al. 2006).

[43] As this article was going to press, one of these papers was published (Williams 2013). Nothing in it leads me to change my assessment of the literature, for I believe the analysis is problematic in a number of important ways. For example, some of the strongest findings result from comparisons between the top two and the bottom two of the six tiers identified in the Bar Passage Study, which is the database that Sander, Williams, and most critics use. But most minority students and almost all blacks in the bottom two tiers are

every social scientist I am aware of who has examined Sander's data and methods rejects the conclusions he has drawn (see, e.g., Ayres & Brooks 2005, Ho 2005, Rothstein & Yoon 2008, Camilli et al. 2011). The definitive statement about the value of Sander's research is, however, found not in the usual scientific literature but in an amicus brief submitted to the Supreme Court (Brief Empir. Schol. 2012) that responds to a similar brief that Sander and his *Mismatch* coauthor Taylor filed in conjunction with the Court's consideration of *Fisher*. Most signers of the brief written in response were new to the affirmative action controversy and had not previously reviewed Sander's efforts. They include leading methodologists in five different fields (economics, law, political science, sociology, and statistics), two of whom (Gary King and Donald Rubin) are members of the National Academy of Science.[44] After reviewing Sander's work, Williams's contributions, and other studies cited by Sander and Taylor in support of the mismatch hypothesis, these experts concluded that

> whether one finds Sander's conclusions highly unlikely or intuitively appealing, his "mismatch" research fails to satisfy the basic standards of good empirical social-science research. The Sander-Taylor Brief misrepresents the acceptance of his hypothesis in the social-science community and, ultimately, the validity of mismatch. Numerous examples exist of better ways to perform the type of research Sander undertook. Sander's failure to set up proper controls to test his hypothesis and his reliance on a number of contradictory assumptions lead him to draw unwarranted causal inferences. At a minimum, these basic research flaws call into question the conclusions of that research.
>
> In light of the many methodological problems with the underlying research, *amici curiae* respectfully request that the Court reject Sander's "mismatch" research. . . . (Brief Empir. Schol. 2012, pp. 27–28)

In promoting his mismatch theory, Professor Sander has at times gone out of his way to denigrate a study that David Chambers, Terry Adams, and I published showing that Michigan's African American and Hispanic law students, most of whom benefitted from affirmative action, did about as well as its white graduates in their post–law school careers (Lempert et al. 2000). He argues that our sample was seriously biased so as to favor those who had achieved success. He particularly disputes our estimate that close to 95% of Michigan's minority graduates pass a bar. His estimate, based on an analysis of how Michigan's African American graduates fared on the bar during a three-year period for which he had bar results, test taker names, and law school photo books from which he could code test taker race, is far lower:[45]

> The results suggested that UMLS [University of Michigan Law School] blacks taking a bar exam for the first time had a 62% pass rate; those taking multiple bar exams had an eventual success rate of 76%. . . . This finding is, in my view, devastating to Lempert's study and to his testimony in *Grutter*. . . . Taking attrition

in tier 6, which is composed of historically and still largely black law schools. Thus, any study that finds that students in the bottom tier perform relatively better, given their admissions credentials, than students in the top two tiers is contaminated by the possibility that students at largely black law schools perform better than expected not because they are less mismatched but because they are at schools with a critical mass of minority students, for other reasons feel more comfortable at these schools than they do at predominantly white law schools, or tend to take bar exams in states where passing is easier. Williams never alerts his readers to the unusual locations of black students in the bottom two tiers or the plausible rival hypothesis that contaminates his results.

[44] I was one of the signers, but the methodological analysis was not mine, and although I concur in the brief's conclusions, I contributed little to its preparation.

[45] Data like these with names attached are confidential and protected by law. Professor Sander acquired the data as part of the discovery process in litigation over a voter-passed initiative that barred Michigan's state schools from employing race-conscious admissions. The relevance of the data acquired to questions at issue is hard to see. Rather, it appears the request was used to provide Professor Sander with confidential information he could have acquired in no other way.

> at the University of Michigan into account, conservatively, the [aggregate graduation and first-time bar passage rate] for black Michigan students during the same period is 60%. This simple comparison thus suggests that the mismatch effect sharply lowers the success rates of the purported beneficiaries of affirmative action at UMLS. (Sander 2011, p. 943)

Sander's results are facially implausible. If only 60% of Michigan's African American students were graduating and passing the bar on the first try, it would have been noticed.[46] An even stronger clue could be found in the data from the Bar Passage Study that Sander in his other work has relied on. These data reveal that about 94% of the African American students in the study's 14 top-tier schools pass the bar. Because Michigan consistently ranks among the nation's top ten law schools and the test scores and grades of those it admits are consistent with this ranking,[47] a finding that its minority students' eventual bar passage rate is 18% below the average of its peers defies credulity.

What is the true picture? I acquired the data Professor Sander used and attempted to replicate his study. In doing so, I had an advantage over him in that I could check with the law school's admissions staff when a student's photo was ambiguous with respect to race. I could also, with the aid of the staff, specify a race for about two-thirds of the 250 students whose pictures were missing from the photo books Sander and I had available. The results are in **Table 1**.

Over the three years for which there are data, 84.8% of the African American students and 98.3% of the white students who took the bar for the first time during these years passed on their first or a later attempt. The estimate of overall African American bar passage is low, however, not only because some may have taken and passed a bar in states not in the sample,[48] but also because it is common for persons who fail the bar on first try to pass it on a later attempt. The truncation of the data set means there is no later attempt information for graduates in the 2006 cohort, and the results for graduates in this cohort bring the average results down. Looking just at those who took the bar for the first time in 2004 or 2005, we see that 91% of the African Americans in this group had passed at least one bar by the end of 2006.[49] As interesting, there is little evidence that affirmative action deflates Michigan's bar passage results. Hispanics, who benefit from affirmative action, did about as well on the bar as whites, who do not, and African Americans, who benefit from affirmative action, did about as well on the bar as Asians, who do not.

What to Do?

When it comes to influencing policy, there is a strong first-mover advantage. The initial eye-catching study that tells a politically attractive tale has every advantage, whether it catches the public eye because of the compelling quality of the underlying science or because of a concerted effort to disseminate

[46] In the first years after Michigan began its affirmative action program, its minority bar passage rate in Michigan was in the range Sander reports, and it was headline news. Moreover, students don't stop comparing notes after they graduate. If the school's minority alumni were failing the bar in large numbers, the school would have heard.

[47] The academic credentials of many admitted African Americans are below and often substantially below the credentials of those of most admitted white students. But this is true of all or almost all of those schools in the Bar Passage Study top tier. Just as schools like Michigan admit the academically strongest white applicants, so, too, they admit the academically strongest minority applicants, and Michigan's minority applicants can be expected to perform as well on the bar as those from an elite law school peer group that includes at least as many schools with applicant pools slightly weaker than Michigan's as it does schools with applicant pools that are slightly stronger.

[48] Bar passage results were available for fewer than half the states, but these were the states where the overwhelming majority of Michigan's law graduates take the bar.

[49] If these students failed, they would have been able to retake the test from two to five times depending on when they first failed. Numerically, this means that of 33 black Michigan graduates who took the bar for the first time during 2004 and 2005 in the states whose results we know, by 2006 at most 3 had not passed a bar, and some of these may have passed the bar in a nonreporting jurisdiction or after the period for which there are data.

Table 1 Bar exam passage rate by ethnicity, 2004–2006

Ethnicity (all years/ 2004–2005)	First-time pass (%)	First-time fail (%)	Ever pass (%)	No pass record (%)	Ever pass, 2004–2005 (%)	No pass record, 2004–2005 (%)
White (753/483)	95.9	4.1	98.3	1.7	99	1
Asian (89/60)	83.1	16.9	86.5	13.5	95	5
Black (46/33)	78.3	21.7	84.8	15.2	90.9	9.1
Latino (39/12)	97.4	2.6	100	—	100	—

its story, regardless of reliability.[50] As with Ehrlich's (1975) death penalty research, the Minneapolis spouse abuse arrest experiment, or the "broken windows" hypothesis, trenchant criticism or additional research may catch up to an originally misleading story, but by the time it does, the world will have changed, new policies may be institutionalized and hard to dislodge, and some minds changed by the initial evidence may be beyond further change. Just as law and social science now has more to offer policy makers because its theories are better grounded and its students can access better tools and data, poorly done research and research whose import is exaggerated or otherwise misunderstood have a greater capacity to mislead policy makers in harmful ways.

[50]It is not only results that appeal to conservatives that enjoy undue publicity and can be oversold. The Minneapolis spouse abuse arrest study was welcomed more by feminists on the Left than by conservatives on the Right. Or consider the High | Scope Perry Preschool Project (Schweinhart et al. 2004). This almost unique, and in many ways outstanding, study initially reported the short-term effects and over time the much longer-term implications of exposing children raised in poverty to a well-designed preschool program. The gains to those in the experimental group were real and substantial, and the study's results helped provide the intellectual underpinnings for the nation's multibillion dollar decision to make Head Start preschool education generally available. But the Perry Preschool study involved only a small number of children in one midwestern city. Although it provides reason to believe that substantial benefits can come from universal preschool, it does not tell us that they will be realized in all places with all program models, nor does it consider the problems inevitably encountered in trying to bring model programs to scale. Thus, it is not surprising that the best evaluation to date of Head Start nationwide shows that unlike the Perry Preschool experience, the benefits of Head Start tend to be short lived, dissipating rapidly after children enter primary schools (US Dep. Health Hum. Serv. 2010). It is also not surprising that the Perry Preschool narrative still generates support for Head Start.

What can the field and those working in it do? For starters, a little modesty is in order. The urge to conclude empirical analyses with implications for policy is strong, particularly when the targeted audience is readers of law reviews. The tendency to overstate implications and understate caveats is common, with recitations of a study's limitations too often reminiscent of the way side effects are recited by rote in drug commercials. Almost never, however, does one study, even against a background of prior research, justify a policy change. This is true no matter how well done the study or how intuitively plausible the story it tells. Indeed, as intuitive plausibility increases, so does the likelihood of exaggerating a study's implications. Consumers of even the best studies should be alert to these limitations and dangers.

Social scientists who do policy-relevant research should carefully delineate what their work does and does not show. They should not hide data weaknesses or reliance on assumptions, and they should specifically note untested plausible rival hypotheses and limitations on external validity, including the uncertainties that are inevitable when bringing pilot programs to scale.

Exhorting individuals to be more modest is, however, far from all that should be done. Institutional changes are more important. Some helpful changes are happening. One is wider data sharing, made easier by the Internet and data repositories, such as ICPSR (**http://www.icpsr.umich.edu**), and encouraged by requirements for data sharing that numbers of journals and such major funders as the NSF and National Institutes of Health (NIH) have instituted. A second is institutionalized meta-analyses and agenda setting. In this regard, the NRC (the research arm of the National Academy of Sciences) and its standing committee on crime and justice have been particularly important to law and social science. NRC volumes on firearms violence (Wellford et al. 2004) and on deterrence and the death penalty (Nagin & Pepper 2012) show the usefulness of their efforts, although many with strongly intuitive or politically motivated policy preferences have not been convinced. Similarly valuable is the Campbell Collaboration,[51] which commissions topic-specific meta-analyses of social science research, including a number that summarize findings in law and social science (see, e.g., Davis et al. 2008, Mitchell et al. 2012, Mazerolle et al. 2013).

Yet if research in law and social science is to better serve policy makers, more needs to be done. If informing policy were all that mattered, law and social science funders could take a lesson from NIH's disease-focused programs and, rather than fund the best law and social science regardless of topic, do more funding along programmatic lines. The National Institute of Justice (NIJ)-funded spouse abuse studies illustrate what can be gained. Because NIJ funded six quasi-replications[52] of Sherman's original experiment, a more complete and complicated picture of the effects of arrest on spouse abuse emerged. It might appear that from a policy perspective, we are not much further advanced because even if the studies taken together indicate that arrest, in the aggregate, appears unlikely to backfire and might in some cases deter, the effects are small enough and vary enough that implications for policy are far from clear. Yet going from a situation in which mandatory or presumptive arrest policies seem unquestionably effective to one in which they appear problematic and greater use of arrest seems unlikely to be key to stopping spouse abuse is an important advance.

The problem is not the confusion caused by the additional experiments but the confusion that remains owing to lack of further research. As important as what the spousal arrest studies showed is what they do not reveal. They do not reveal the mechanism behind any reduction in spouse abuse that may accompany arrest. Deterrence is assumed to be the mechanism, and although the data are consistent with deterrence theory, they by no means prove it. In the Sherman & Berk (1984) study, for example, much of the reduction in recidivist spouse abuse may stem from the tendency of arrest to break up relationships. Yet if an abuser leaves a woman following an arrest but soon takes up with and abuses another woman, although recidivist abuse experimentally measured will have diminished, the community-wide incidence of spouse abuse may be unchanged and the number of women victimized may increase. The possibility of interactions between treatments and offender and victim characteristics is also largely unexplored, although Sherman's Milwaukee experiment suggests that whether arrest helps protect women or further endangers them depends on victim and offender characteristics (Sherman et al. 1992a). Finally, a policy maker would want to know about costs and benefits of arrest for misdemeanor spouse abuse beyond repeat violence, such as the implications of arrest for family stability, job retention, and future employability. The spouse abuse replications answer none of these questions.

Imagine that a proposed cancer treatment were being evaluated, and a first study found that as compared to controls, a statistically significant proportion of mice given the treatment saw their cancers go into remission. Next, the drug might be tested in different doses in

[51] The effort is named in honor of Donald Campbell, who touched many more lives than mine with his teaching and scholarship.

[52] I use the term quasi-replications because even though each replication was required to randomize the arrest treatment and measure effects by both police records and spousal interviews over a period of at least six months, other aspects of these experiments, such as what followed upon arrest, differed across the research sites and from the original Minneapolis study. Although six follow-up studies were funded, the one in Atlanta was not completed, so the published literature includes the results from only five of the studies, those done in Charlotte, Colorado Springs, Dade County, Milwaukee, and Omaha.

different animal models. If the results were still positive, side effects would be noted, studies might be undertaken to see how they could be ameliorated, and genetic analyses might be done to see if the drug was effective not phenotypically but only with animals having certain genotypes. Further positive outcomes might justify human trials, first to ensure the drug's safety and then to establish its effectiveness. Even if the drug were approved for use, monitoring would continue to update what had been learned. Looking back, we might see that tens of millions of dollars had been spent to test and develop a drug that was only partially effective against only one type of cancer and then only if a sufferer had a certain genotype. Most likely we would think the expenditure well worth it.

Ideally, we would take the same programmatic approach to dealing with social pathologies. But we seldom do. To the extent we accumulate relevant bodies of policy-deployable knowledge, this is less likely to come from a programmatic effort to collect relevant data, explore complex interactions, and account for the myriad plausible hypotheses that complicate scientific interpretation than it is to be the outcome of a series of separately funded investigations that have been selected for support on their own scientific merits with little regard to whether they are the next step in growing a body of policy-related knowledge. A too common result is a body of policy-relevant research with important gaps that in the best case are filled in by well-grounded theory, in the intermediate case by what makes the most intuitive sense, and in the worst case by wishful thinking or a failure to recognize that gaps exist. Although research in law and social science has much to offer policy makers and will have more in coming years, the lack of programmatic funding and thinking has yielded too many examples of studies that have proved unduly influential because of their political rather than their scientific appeal.

Yet I cannot be a full-throated advocate for programmatic research in law and social science. The reason is simple—money. If effective policy requires programmatic efforts, scientific advances require research that breaks new paths, is often idiosyncratic, and initially has no clear policy relevance. One could make a strong case that our country should invest adequately in both programmatic and investigator-initiated efforts, but there is no chance that the social sciences, much less that small corner of the social sciences we call law and social science, will receive NIH Cancer Institute–level funding. Indeed, the aggregate of law and social science funding from all sources, including NSF, NIJ, and private foundations, does not exceed and is unlikely ever to exceed a fraction of the funding that within NIH is devoted to specific diseases and conditions. The danger in advocating programmatic approaches to help resolve policy conundrums is not just that basic-science funding will be slighted but also that research on a number of important problems will receive no funding whatsoever. It is no surprise that much of the scientifically strongest policy-relevant research within law and social science has been done by criminologists with a focus on policing. Funding has been available because reducing crime is a universally approved, politically popular goal that often carries with it a sense of urgency. The civil justice system, by contrast, has huge economic impacts and important implications for power and equality, but relatively little research money has been directed its way in recent years. Nor is there a civil justice equivalent of the coordinated experiments that were undertaken to understand what works in bail setting, arrest for spouse abuse, and hot spot policing.

Equally unfortunate would be funding policy-relevant research to the exclusion of work that seeks to elucidate the law and society relationship with no goal other than an improved understanding of some aspect of the law or legal system. Some of the most interesting and important contributions to law and social science are of this type. My favorites include Kitty Calavita's (1992) effort to understand the social dynamics of US immigration policy, Stuart Scheingold's (1974) and Michael McCann's (1994) studies of the role that law and rights consciousness can play in progressive social movements, Bob Ellickson's (1991) study of how disputes involving wandering cattle are resolved in Shasta County, California, and the many articles and occasional books that have contributed to what is now a vast store of knowledge about jurors and the jury system (e.g., Vidmar & Hans 2007).

Getting at Mechanism

Reflecting on this list of favorites, I am struck by the fact that with one exception, the books that first came to mind rely largely or exclusively on qualitative methods, but the studies I cited for their influence on policy, whether as good or bad examples, are quantitative in their approach. Qualitative research can be rigorous, but there is a rigor to quantitative methods, in apparent precision if nothing else, that qualitative research lacks. Moreover, although the subjectivity inherent in quantitative research is often underappreciated and too often disguised by the methods employed, the subjective element in qualitative research not only is more transparent but in the usual case bears a stronger relation to what is reported if for no other reason than that it is harder for a qualitative researcher to test fairly the implications of competing theories. Quantitative analyses are also more amenable to checking and replication because in the usual case quantitative data exist apart from the researcher's interpretation.[53]

Qualitative approaches are often better than quantitative ones at getting at mechanism, the often complex processes that mediate between conditions and outcomes. I think it is for this reason that I admire the works I mention. The one book based mainly on quantitative studies, the Vidmar & Hans (2007) volume on the jury, is the exception that proves the rule. It incorporates so many studies on so many aspects of jury behavior that we leave it not only with a good understanding of how juries behave but also with a good understanding of why they behave as they do.

A failure to understand mechanism can lead to mistaken conclusions about reasons for behavior, and the result can be poor policy decisions. I discovered this in examining the decisions of a public housing eviction board in Honolulu, Hawaii (Lempert & Monsma 1994). I had reason to expect that Samoans were discriminated against in board decisions. They occupied the lowest rung of Hawaii's class structure and were most likely to be stereotyped as violent and shiftless. A quantitative analysis confirmed my expectation, for Samoans owing rent were more likely than tenants of other ethnicities to be evicted, holding all other measured variables equal. But the solution of more actively enforcing antidiscrimination laws would not have helped Samoan tenants. Talking to board members, project managers, and others revealed that the Samoans suffered most from the cultural unacceptability of the excuses they gave, and it was hard to label the rejection of their excuses discrimination.[54] Moreover, increased efficiency in processing evictions hit Samoans harder than others because although they were more likely to have kin networks they could mobilize to help with the rent, this could take some time.

Inattention to mechanism is a failing common to several of the studies I have criticized as

[53] Most quantitative data are derived from such sources as opinion surveys, government records, experimental outcomes, physiological measures, and the like. What these measures have in common is that they are produced separately from the researcher's interpretation of what they mean and are usually available to other investigators in forms that are unmediated by the person who initially secured them. This does not mean that the researcher's subjectivity has not shaped the data. Different ways of asking for the same information will yield different replies, and a data set will include only that information that the person who assembled it thought to acquire. As noted earlier, however, as technology allows qualitative researchers to record and share the raw data they are examining, some of the distinctions between the data that feed qualitative and quantitative analyses are breaking down.

[54] "My child fell ill and I needed my rent money to pay the doctor" is an excuse that might leave a board member willing to give a tenant in default more time to pay back rent. But the excuse "I sent my rent money to Samoa to help pay the cost of my uncle's funeral" would not lead a board member to advocate leniency. Yet in traditional Samoan culture, the latter excuse is every bit as compelling as the ill child excuse is in modern American culture.

unduly influential. Ironically, it may be fostered by the desire to play to the policy audience—if something works, policy makers may not care why it works.[55] But they should. Sound policy often requires knowledge of mechanism. If those arrested when police interrupt an incident of spouse abuse become more fearful of a police call and so are less likely to recidivate than those not arrested, then a presumptive arrest policy is likely to be good policy. If they are less likely to abuse their spouses because they have left them and are living with and abusing other women, then arrest has merely changed the identity of those victimized. If arrest deters future spouse abuse by men whose social position means they have a lot to lose but leads men with little to lose to take violent revenge on spouses who have called the police, then a presumptive arrest policy may aid some women while harming others with violence greater than the violence prevented. Without knowing more about mechanism, we cannot say whether presumptive arrest is a wise policy, nor can we tailor responses to offender and victim characteristics.

The works by Ehrlich (1975) and by Lott & Mustard (1997) show a different hazard to ignoring mechanism: the acceptance of results that don't hold up to close scrutiny. Ehrlich and Lott and Mustard believe their findings show deterrence at work, but understanding what deterrence theory implies should make policy makers less rather than more willing to accept the claims these authors advance. In the case of Ehrlich, if deterrence is the mechanism by which the death penalty works to reduce crime, then one would expect to find a deterrent effect to the death penalty if the time series relating executions to homicides terminates before the early 1960s. When one doesn't, and in some specifications even sees the direction of effects reverse, there is good reason to doubt that the death penalty deters killing. As for Lott and Mustard, if knowing people might be carrying concealed weapons deters killing, one would not expect homicide rates to rise and fall dramatically with drug wars because those on both sides know that those on the other side carry guns. Moreover, one would not expect right-to-carry laws to be associated with an increase in property crimes because those carrying guns might be expected to use them if their property were being taken.

The articles by Wilson & Kelling (1982) and by Sander (2004) both attend to mechanism and show another virtue of such attention: It facilitates tests of a theory and allows for its falsification. Because Wilson and Kelling claimed that signs of neighborhood disorder such as broken windows and people loitering on the corner promoted crime, Sampson and Roundenbush could call the theory into question by carefully recording visible signs of disorder and showing that after controlling for relevant variables, visible disorder did not correlate with neighborhood crime rates.[56] Because Sander identifies mismatch between the admissions credentials of minority law school students and the average credentials of the

[55] Perhaps even more ironically, insufficient attention to mechanism may arise in quantitative research because the researcher is following the classic hypothesis-testing paradigm in order to test mechanism. Thus, an experimenter interested in whether arrest has a deterrent effect on future spouse abuse will hypothesize that it does, and if those arrested are less likely to recidivate than the controls, the experimenter will conclude that the deterrence hypothesis is supported and that deterrence is the likely mechanism by which arrest works. The experimenter's mistake is that although the results make the deterrence hypothesis more likely, it does not follow that deterrence is the likely mechanism by which arrest works. Other possible mechanisms may not have been made less likely by the support the deterrence hypothesis received, and a close look at the data may reveal that some rival hypotheses receive more support than the deterrence hypothesis that the experimenter, to the exclusion of other possible mechanisms, focused on.

[56] "Broken windows" policing has come to stand for more than repairing broken windows and in other ways making neighborhoods look safer and less neglected. In New York City, which is the poster child for the supposed success of this approach, and in some other cities, the concept has been operationalized as a more aggressive approach to neighborhood policing that involves an increased police presence, frequent stops and searches of those encountered on the streets, and more arrests for minor crimes. These activities may reduce crime even if sprucing up neighborhoods does not. They also impose costs on innocent people that are quite different from the repair and clean-up costs that removing signs of disorder calls to mind.

schools they attended as the mechanism behind the disproportionately low graduation rates and high bar-failure rates of African American law students, his claims could be called into question by showing that controlling for admissions credentials, African American students, unless they were at historically black law schools, tended to do better on both measures if they were in higher- rather than lower-tier law schools, with those in the top tier having the greatest success (Chambers et al. 2005).

IN CONCLUSION: A BOW TO DON CAMPBELL

I conclude this essay, almost where I began: at the start of my life as a social scientist and the seminar I took with Don Campbell. The seminar was focused on methods but had almost nothing to do with statistics. Campbell was not hostile to sophisticated statistics, far from it, but the usefulness of statistics depended for him on sound methods, as did the right choice of what statistics to use. Campbell never confused statistical with methodological sophistication. A major theme of the seminar was the logic of causation and the need to understand causal logic in the context of the problem to be investigated. Only with such understanding can one design a study that controls as well as possible for plausible threats to a desired causal inference. Similar understanding is needed after the results are in to recognize threats that remain and what causal claims can be safely made.

Campbell's shorthand summary of what good research design entailed was "the control of plausible rival hypotheses." Any study aimed at establishing a causal relationship should as far as possible be designed so as to eliminate plausible competing claims about a relationship's causes if the competing claims are false.[57] For example, suppose that following a year in which traffic deaths were at an all-time high, a state decided to crack down on speeding by passing a law that punished speeders more severely. If the year after the law was passed, traffic deaths dropped substantially, this could be due to the law, but it could also result from regression to the mean (another concept whose implications Campbell emphasized). A study that compared the number of traffic deaths the year before the crackdown with the number of deaths in the following year could not distinguish between the law's effects and falloff due to statistical regression. If, however, the study examined data for several years before and after the crackdown, then the possibility that the falloff was regression to the mean could be distinguished from the likelihood that the crackdown had had an effect (Campbell & Ross 1968).

In his formulation of what research designs should aim to control, Campbell emphasized the word plausible. A creative mind can construct a story to explain almost anything, and lawyers seeking to deny the implications of social science evidence are adept at doing so. But many possible stories are not plausible. If a possible explanation is not plausible, failure to control for it provides little, if any, reason to question a study's conclusions. Similarly, when evaluating proposed research, designs should not be faulted because they leave implausible possibilities open. Plausibility also plays an important role in weighing purported findings. It is not, for example, surprising that Professor Sander's claim that only 76% of Michigan's African American law school graduates passed the bar did not stand up to replication, for there is no plausible reason why one elite school's African American graduates would do so poorly on the bar when 94% of the African American graduates of other elite schools succeeded in passing. Similarly, the Lott and Mustard research is questionable on its face because although one can suggest that substitution effects explain why the same law that diminishes homicides and rapes increases larceny and auto theft, no accepted understanding of motivations to crime makes the substitution hypothesis plausible.

Another concern that Campbell drove home was the importance of attending to how key

[57]For an equally instructive and somewhat different statement of the same point, see Arthur Stinchcombe's (1968) marvelous book *Constructing Social Theories*.

concepts are operationalized. When data are analyzed, the analyst seldom manipulates the concept of interest. What is operated on in data analysis is some manifestation of the concept, and statements that seemingly refer to the concept in fact refer not to some abstract meaning but to how the concept is manifested in the data, or its "operationalization." It is easy to lose sight of this and to think in terms of the concept rather than what has actually been measured. Thus IQ, which is a score on a test, is often treated as if it is intelligence. But IQ at best measures only certain aspects of intelligence, and scores assigned to individuals may be contaminated by lucky or unlucky guessing, fatigue, and test familiarity, among other confounds. Policy makers should look carefully at how concepts are operationalized before relying on a study's results to guide policy, and researchers who seek to affect policy should, before deciding how to operationalize their variables, think about what information policy makers most need. For example, when Maxwell et al. (2001) combined for reanalysis the data from the five spouse abuse study replications, they treated reports of threatened aggression by those they interviewed as a form of recidivist abuse. Some might argue that unconsummated threats do not constitute criminal spouse abuse and that preventing threats that go nowhere is a lesser concern of the police than preventing violence, if it is a concern at all. This decision about how to operationalize spouse abuse, which is easy to overlook or forget when reading the synthesis, may explain why statistically significant effects of arrest are found only in the interview data and not in the police records. It is possible that if the operationalization of abuse in the global study excluded verbal aggression, the interviews would show no diminution in violence due to arrest, and they might even show an increase in harm. I expect that police departments deliberating on the wisdom of establishing presumptive arrest policies would be more interested in the effects of arrest on later violence than in its effects on an undifferentiated measure of threatened and actual assaults.

Campbell also pointed out that when one left tabular displays for more complicated statistics, one inevitably lost information even if one gained insights and capacity for control. The lesson here is to look closely at data and to recognize that the story they tell can change with how they are examined. When a more complex analysis changes what a simpler analysis seems to reveal, this does not mean that the results of the more complex analysis are suspect. It does mean that one should seek to determine why two (or more) ways of looking at the same data present different pictures. Too often it may be because a researcher's choices tipped the scales in the direction of a favored perspective. This is a particular problem because tipping the scales is relatively easy and does not require dishonesty. If one explores different model specifications, random error may mean that a particular specification exaggerates the likely importance of a crucial policy-relevant variable. If only the model showing significance is presented, those relying on it will be misled. At a minimum, reported significance levels will be unreliable because the more opportunities a variable has to achieve statistical significance, the greater the likelihood that it will be found significant on at least one occasion, even if chance is responsible for the association. Yet researchers can be strongly tempted to search for specifications that make crucial variables significant even when they have no particular axe to grind because it is most often easier to publish positive findings than negative ones.[58]

[58]This gives rise to the so-called file drawer problem. If 19 researchers do a study and do not achieve significant results, the article they intended to write remains in the file drawer. If the twentieth researcher doing the same study achieves significant results, the article will be submitted to a journal and published. This no doubt exaggerates what occurs, but it contains an important truth. To the extent there is a bias in favor of publishing positive results, relationships labeled significant may not be as unlikely as they appear because if every study that examined the same relationship were known, the chance that the relationship would be significant in one study would be greater, and perhaps far greater, than the published significance statistics suggest.

This is one reason why I argue that empirically driven policy setting should require a body of consistent research and should never be driven by a study or two. If different studies don't agree, the inconsistency should be taken seriously. Policy makers should not automatically prefer what they see as the more statistically sophisticated research.[59] Rather, the sources of inconsistency should be identified and judgments made on that basis.

I refer to these methodological principles and others that I took from Campbell as soft methods because they do not require the mastery of mathematics that expertise in statistics entails. But there is nothing else soft about them. They require rigorous thinking and close attention to the details of the research enterprise.[60] Although they are accessible to almost every researcher working with empirical data, too often their implications are ignored. If the goal is to responsibly inform public policy through empirical research, those doing research in law and social science must learn the lessons Don Campbell taught, whether or not they were lucky enough to have studied with him.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Abel RL. 1974. A comparative theory of dispute institutions in society. *Law Soc. Rev.* 8:217–347

Ayres I. 1991. Fair driving: gender and race discrimination in retail car negotiations. *Harvard Law Rev.* 104(4):817–72

Ayres I, Brooks RRW. 2005. Does affirmative action reduce the number of black lawyers? *Stanford Law Rev.* 57:1807–54

Ayres I, Donohue JJ III. 2003. Shooting down the more guns, less crime hypothesis. *Stanford Law Rev.* 55:1193–312

Ballew v. Georgia, 435 U.S. 223 (1978)

Beny LN. 2008. Do investors in controlled firms value insider trading laws? International evidence. *J. Law Econ. Policy* 4(2):267–310

Black D. 1979. *The Behavior of Law*. New York: Academic

Boruch R, Snyder B, DeMoya D. 2000. The importance of randomized field trials. *Crime Delinq.* 46:156–80

Brady HE, Collier D, eds. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman & Littlefield. 2nd ed.

Braga AA, Weisburd DL. 2010. *Policing Problem Places: Crime Hot Spots and Effective Prevention*. New York: Oxford Univ. Press

Braga AA, Weisburd DL, Waring EJ, Green Mazerolle L, Spelman W, Gajewski F. 1999. Problem-oriented policing in violent crime places: a randomized controlled experiment. *Criminology* 37(3):541–80

Brief Empir. Schol. (Empirical Scholars as Amici Curiae in Support of Respondents). 2012. See *Fisher v. Texas*. **http://www.utexas.edu/vp/irla/Documents/ACR%20Empirical%20Scholars.pdf**

Calavita K. 1992. *Inside the State: The Bracero Program, Immigration, and the I.N.S.* New York: Routledge

[59]References to sophisticated methods are often seen as referring to methods on the cutting edge of modern statistics or to methods of considerable mathematical complexity. Sophistication in choosing a statistical approach depends, however, on which analytic approaches are best suited to the problem at hand. Sometimes simpler methods will be better.

[60]The same may be said of statistical thinking. The thinking of the hard methodologist and that of the soft methodologist have much in common, and many of the best social scientists are masters of both. My respect for statistical thinking is such that when my daughter went to college, I told her that I did not care what she studied so long as she took at least one course in statistics. Somewhat to my surprise and greatly to my satisfaction, she did.

Camilli G, Jackson D, Chiu C-Y, Gallagher A. 2011. The mismatch hypotheses in law school admissions. *Widener J. Law Econ. Race* 2:165–209

Campbell DT, Ross HL. 1968. The Connecticut crackdown on speeding: time-series data in quasi-experimental analysis. *Law Soc. Rev.* 3:33–53

Chambers DL, Clydesdale TT, Kidder WC, Lempert R. 2005. The real impact of eliminating affirmative action in American law schools: an empirical critique of Richard Sander's study. *Stanford Law Rev.* 57:1855–98

Danziger S, Gottschalk P. 1985. Social programs—a partial solution to, but not a cause of poverty: an alternative to Charles Murray's view. In *Losing Ground: A Critique*, ed. S McLanahan, G Cain, M Olneck, I Piliavin, S Danziger, P Gottschalk, pp. 73–91. Madison, WI: Inst. Poverty Res.

Davis RC, Weisburd D, Taylor B. 2008. Effects of second responder programs on repeat incidents of family abuse: a systematic review. *Campbell Syst. Rev.* 4(15)

Diamond SS, Vidmar N, Rose M, Ellis L, Murphy B. 2003. Jury discussions during civil trials: studying an Arizona innovation. *Univ. Ariz. Law Rev.* 45:1–81

Ehrlich I. 1975. The deterrent effect of capital punishment: a question of life and death. *Am. Econ. Rev.* 65:397–417

Ellickson RC. 1991. *Order Without Law: How Neighbors Settle Disputes*. Cambridge, MA: Harvard Univ. Press

Felstiner WLF. 1974. Influences of social organization on dispute processing. *Law Soc. Rev.* 9:63–94

Fisher v. Texas, No. 11–345, slip op. (U.S. June 24, 2013)

Gendreau P, Ross R. 1987. Revivification of rehabilitation: evidence from the 1980s. *Justice Q.* 4:349–408

Gregg v. Georgia, 428 U.S. 153 (1976)

Grutter v. Bollinger, 539 U.S. 306 (2003)

Ho DE. 2005. Why affirmative action does not cause black students to fail the bar: a reply to Sander. *Yale Law J.* 114:1997–2004

Jencks C. 1985. *How Poor Are the Poor?* New York: Rev. Books

Kalven H Jr. 1968. The quest for the middle range: empirical inquiry and legal policy. In *Law in a Changing America*, ed. GC Hazard Jr, pp. 56–74. Englewood Cliffs, NJ: Prentice-Hall

Kalven H Jr, Zeisel H. 1966. *The American Jury*. Boston: Little, Brown

Katz J. 1988. *Seductions Of Crime: Moral and Sensual Attractions in Doing Evil*. New York: Basic Books

Kornstein D. 1976. A Bayesian model of harmless error. *J. Legal Stud.* 5:121

Kronman A. 1983. *Max Weber*. Stanford, CA: Stanford Univ. Press

Lempert R. 1966. Strategies of research design in the legal impact study: the control of plausible rival hypotheses. *Law Soc. Rev.* 1(1):111–32

Lempert R. 1972a. Norm-making in social exchange: a contract law model. *Law Soc. Rev.* 7:1–32

Lempert R. 1972b. Law school grading: an experiment with pass-fail. *J. Legal Educ.* 24:251–308

Lempert R. 1975. Uncovering "non-discernible" differences: empirical research and the jury-size case. *Mich. Law Rev.* 73:643–708

Lempert R. 1977. Modeling relevance. *Mich. Law Rev.* 75:1021–57

Lempert R. 1981. Desert and deterrence: an assessment of the moral bases of the case for capital punishment. *Mich. Law Rev.* 79:1177–231

Lempert R. 1989. Humility is a virtue: on the publicization of policy-relevant research. *Law Soc. Rev.* 23:145–61

Lempert R. 2010. The inevitability of theory. *Calif. Law Rev.* 98:877–906

Lempert R, Chambers DL, Adams TK. 2000. Michigan's minority graduates in practice: the river runs through law school. *Law Soc. Inq.* 25:395–505

Lempert R, Ikeda K. 1970. Evictions from public housing: effects of independent review. *Am. Sociol. Rev.* 35:852–59

Lempert R, Kidder WC, Clydesdale TT, Chambers DL. 2006. *Affirmative action in American law schools: a critical response to Richard Sander's "A Reply to Critics."* Work. Pap. 60, Prog. Law Econ. Arch.: 2003–2009, Univ. Mich., Ann Arbor. **http://law.bepress.com/umichlwps-olin/art60**

Lempert R, Monsma K. 1994. Cultural differences and discrimination: Samoans before a public housing eviction board. *Am. Sociol. Rev.* 59:890–910

Lempert R, Saltzburg SA. 1977. *A Modern Approach to Evidence*. St. Paul, MN: West

Lott JR, Mustard DB. 1997. Crime, deterrence, and right-to-carry concealed handguns. *J. Legal Stud.* 26:1–68
Martinson R. 1974. What works?—Questions and answers about prison reform. *Public Interest* (35):22–54
Martinson R. 1979. New findings, new views: a note of caution regarding sentencing reform. *Hofstra Law Rev.* 7:242–58
Maxwell CD, Garner JH, Fagin JA. 2001. *The effects of arrest on intimate partner violence: new evidence from the Spousal Assault Replication Program*. Res. Brief, Natl. Inst. Justice, Washington, DC
Maxwell CD, Garner JH, Fagin JA. 2002. The preventive effects of arrest on intimate partner violence: research, policy and theory. *Criminol. Public Policy* 2:51–80
Mayhew L. 1968. *Law and Equal Opportunity: A Study of the Massachusetts Commission Against Discrimination*. Cambridge, MA: Harvard Univ. Press
Mazerolle L, Bennett S, Davis J, Sargeant E, Manning M. 2013. Legitimacy in policing: a systematic review. *Campbell Syst. Rev.* 9(1)
McCann M. 1994. *Rights at Work: Pay Equity Reform and the Politics of Legal Mobilization*. Chicago: Univ. Chicago Press
Mitchell O, Wilson D, Eggers A, MacKenzie D. 2012. Drug courts' effects on criminal offending for juveniles and adults. *Campbell Syst. Rev.* 8(4)
Murray C. 1984. *Losing Ground: American Social Policy, 1950–1980*. New York: Basic Books
Nagin DS, Pepper JV, eds. 2012. *Deterrence and the Death Penalty*. Washington, DC: Natl. Acad. Press
Normand J, Lempert R, O'Brien CP, eds. 1994. *Under the Influence? Drugs and the American Work Force*. Washington, DC: Natl. Acad. Press
O'Gorman HJ. 1963. *Lawyers and Matrimonial Cases: A Study of Informal Pressures in Private Professional Practice*. Glencoe, IL: Free Press
Pager D, Western B, Bonikowski B. 2009. Discrimination in a low-wage labor market: a field experiment. *Am. Sociol. Rev.* 74:777–99
Regents of University of California v. Bakke, 438 U.S. 265 (1978)
Rehavi MM, Starr SB. 2012. *Racial disparity in federal criminal charging and its sentencing consequences*. Pap. 12–002, Law & Econ., Empir. Legal Stud. Cent., Univ. Mich., Ann Arbor. **http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1985377**
Rothstein J, Yoon AH. 2008. Affirmative action in law school admissions: What do racial preferences do? *Univ. Chicago Law Rev.* 75:649–714
Sampson RJ, Raudenbush SW. 1999. Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am. J. Sociol.* 105:603–51
Sampson RJ, Raudenbush SW. 2004. Seeing disorder: neighborhood stigma and the social construction of "broken windows." *Soc. Psychol. Q.* 67(4):319–42
Sander RH. 2004. A systemic analysis of affirmative action in American law schools. *Stanford Law Rev.* 57:367–483
Sander RH. 2005. A reply to critics. *Stanford Law Rev.* 57:1963–2016
Sander RH. 2011. Listening to the debate on reforming law school admissions preferences. *Denver Law Rev.* 88:889–953
Sander RH, Taylor S. 2012. *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit It*. New York: Basic Books
Sarat A, Felstiner WL. 1995. *Divorce Lawyers and Their Clients: Power and Meaning in the Legal Process*. New York: Oxford Univ. Press
Scheingold SA. 1974. *The Politics of Rights: Lawyers, Public Policy, and Political Change*. New Haven, CT: Yale Univ. Press
Schweinhart LJ, Montie J, Xiang Z, Barnett WS, Belfield CR, Nores M. 2004. *Lifetime Effects: The High|Scope Perry Preschool Study Through Age 40*. Ypsilanti, MI: High|Scope Press
Sherman LW, Berk RA. 1984. The specific deterrent effects of arrest for domestic assault. *Am. Sociol. Rev.* 49:261–72
Sherman LW, Cohn EG. 1989. The impact of research on legal policy: the Minneapolis Domestic Violence Experiment. *Law Soc. Rev.* 23:117–44
Sherman LW, Schmidt JD, Rogan DP, Smith DA, Gartin PR, et al. 1992a. The variable effects of arrest on crime control: the Milwaukee Domestic Violence Experiment. *J. Crim. Law Criminol.* 83:137–69

Sherman LW, Smith DA, Schmidt JD, Rogan DP. 1992b. Crime, punishment, and stake in conformity: legal and informal control of domestic violence. *Am. Sociol. Rev.* 57:680–90
Stinchcombe A. 1968. *Constructing Social Theories*. New York: Harcourt, Brace & World
Thomas WH Jr. 1976. *Bail Reform in America*. Berkeley: Univ. Calif. Press
Tribe L. 1971. Trial by mathematics: precision and ritual in the legal process. *Harvard Law Rev.* 84:1329–93
Trubek D. 1972. Max Weber on law and the rise of capitalism. *Wis. Law Rev.* 720–53
US Congress. 1955. *Recording of Jury Deliberations: Hearings Before the Subcommittee to Investigate the Administration of the Internal Security Act of the Senate Committee on the Judiciary*. 84th Cong., 1st Sess., 63–81
US Dep. Health Hum. Serv. 2010. *Head Start impact study.* Final Rep., Adm. Child. Fam., Washington, DC
Vidmar N, Hans VP. 2007. *American Juries: The Verdict*. Amherst , NY: Prometheus
Wellford CF, Pepper JV, Petrie CV, eds. 2004. *Firearms and Violence: A Critical Review*. Washington, DC: Natl. Acad. Press
Wells G, Penrod S. 2011. Eyewitness identification research: strengths and weaknesses of alternative methods. In *Research Methods in Forensic Psychology*, ed. B Rosenfeld, SD Penrod, pp. 237–56. Hoboken, NJ: John Wiley & Sons
Wightman L. 1998. *LSAC National Longitudinal Bar Passage Study*. Newtown, PA: Law Sch. Admiss. Counc.
Williams D. 2013. Do racial preferences affect minority learning in law schools? *J. Empir. Legal Stud.* 10:171–95
Williams v. Florida, 399 U.S. 78 (1970)
Wilson JQ, Kelling GL. 1982. Broken windows: the police and neighborhood safety. *Atl. Mon.* 249:29–38
Zeisel H, Diamond SS. 1974. "Convincing empirical evidence" on the six member jury. *Univ. Chicago Law Rev.* 41:281–95