

University of Michigan Law School

University of Michigan Law School Scholarship Repository

Articles

Faculty Scholarship

2008

Empirical Research for Public Policy: With Examples from Family Law

Richard O. Lempert

University of Michigan Law School, rlempert@umich.edu

Available at: <https://repository.law.umich.edu/articles/2399>

Follow this and additional works at: <https://repository.law.umich.edu/articles>



Part of the [Legal Writing and Research Commons](#), and the [Public Policy Commons](#)

Recommended Citation

Lempert, Richard O. "Empirical Research for Public Policy: With Examples from Family Law." *Journal of Empirical Legal Studies* 5, no. 4 (2008): 907-926.

This Article is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

EMPIRICAL RESEARCH FOR PUBLIC POLICY: WITH EXAMPLES FROM FAMILY LAW AND ADVICE ON SECURING FUNDING*

Richard Lempert

Perhaps more than in any other field, legal scholarship has aimed directly at influencing public policy. Hence, it is not surprising that empirical scholarship on law related issues often seems to have an agenda that extends beyond the common social science goals of adding to our knowledge base and understanding of human behavior to suggesting to policy makers and practitioners legal and administrative changes that will ameliorate problems they confront and, by the researcher's lights, make this a better world in which to live.

There are, however, many dangers in relying on empirical work, rather than normative arguments, to influence policy.¹ Empirical work is social science, and the standards of social science as well as what it determines can change. Methods, data availability and, to some extent, data quality are improving all the time. What was state of the art research yesterday may be perceived as hopelessly flawed tomorrow, and reanalyzing even the same data using more modern methods may reveal that the earlier advice was not just unsupported by the data but contrary to the policies that the data, properly analyzed, suggest. Institutionalized procedures are slow to change, however, even when their rationales have disappeared, and the law can be even slower, particularly if interest groups have benefit from the status quo regardless of whether it is in the public interest. Thus no matter how much an empirical scientist likes the results of his or her own research, a certain humility about advocating policy changes based on it is in order.

Yet policy makers must act. Agencies should strive to be more effective, and legislators want to enact those laws most likely to advance well being. These bodies always act with some conception of the status quo and the dynamics of human behavior in mind. Even if today's social science findings may be withdrawn tomorrow, if the science is done well they are

* Draft Version, please do not cite or quote without permission.

¹ The classic critique of doing so is Herbert Wechsler's criticism of the use that Chief Justice Earl Warren made of Dr. Kenneth Clark's doll studies in *Brown v. Board of Education*. Toward Neutral Principles of Constitutional Law, 73 Harv. L. Rev. 1 (1959).

the best current estimate of the state of the relevant world and of how people and organizations are likely to respond to certain actions. In particular social science research can discipline the wishful thinking that can lead people to ignore problems they don't want to confront or believe that the policies that most help them will be best for everyone or that policies they disapprove of do no good. It can also reveal complexities that mean no single solution will work everywhere or that solving one problem will raise problems in other areas. Thus research on police highway stops can reveal whether unproductive racial profiling is occurring. Research on bilingual versus English only education can tell us which approach is more conducive to English learning and or other subject learning among non-native English speakers as well as whether effects vary by student characteristics. Research on bankruptcies can tell us something about the role credit card debt, job loss and health problems in pushing people over the edge.

Sound research on matters like these can help policy makers and administrators devise more effective laws and procedures. But empirical studies are almost never a magic bullet. In a society where fifty years of research was necessary to convince almost everyone who matters that smoking is a cause of lung cancer and heart disease and where some are still not convinced of the soundness of climate change science or the likelihood of evolution, one cannot expect that a few empirical studies will resolve issues that have divided the public. Nor should they, for in many areas empirical studies will produce conflicting results with conflicting implications for policy. Work on capital punishment is a prominent current example.

There is also the issue of mechanism and its close relative, theory. Sometimes it is sufficient to know that something works, and we need not worry (too much) why. Aspirin cured headaches long before people knew why it helped, and those with aching heads would have been silly had they refused aspirin until they knew the mechanism that led to their relief. But in the case of policy relevant empirical research, knowing mechanism is generally quite important. This is in part because people and places differ. What works with one group or in one place may not work with another group or in another place. Also knowing mechanism allows for better targeted and more efficient interventions. Thus if a preschool program increased reading readiness because of the phonics approach it employed, it would suggest different kinds of preschool interventions than if the key were that each child received one on one attention from a student volunteer. Theory which posits the mechanism(s) by which effects occur has the added virtue of allowing strong tests of what we think data reveal. Thus if we

theorize that phonics is the mechanism by which a pre-school reading program works to improve performance, we can experimentally try phonics and alternative approaches to reading while hold one on one time with volunteers constant. If the phonics group does better we have greater reason to believe that it is the (or a) key to the program's success. Since the world is complex and outcome soften have multiple causes, considerable research may be necessary to develop a theory complete enough to provide a reliable guide to policy.

A story is told of an old country rabbi who was asked by a disciple what he thought was the best cut of meat. The rabbi told the student it was tongue. The disciple then asked the rabbi what in his opinion was the worst cut of meat. The rabbi again said, "tongue." "How can tongue be the worst cut of meat when it is the best cut of meat?" said the student confused as well as a little indignant thinking he was being trifled with. "There is," said the rabbi, "nothing so good as a good tongue and nothing so bad as a bad one." So it is with empirical work and public policy. Nothing is so helpful as good empirical research and nothing can be so bad as poor research that becomes influential. But what makes empirical work good or bad? It is to this issue, I shall now turn – drawing my examples from three studies designed to inform policy in the area of family law.

These studies are, first, the work of David Chambers on the enforcement of child support obligations, as presented in his book *Making Fathers Pay*,² second the study of mandatory arrest for misdemeanor spouse abuse as described by Larry Sherman and Richard Berk in an article in the *American Sociological Review*,³ and third the study of the implication of California's no fault divorce law on the well being of male and female divorced spouses as presented by Lenore Weitzman in *The Divorce Revolution*⁴.

Chambers: Making Fathers Pay

Let me begin with Professor Chamber's work which, since David Chambers is a long time colleague and friend, I know best. It, moreover, should be a heartening tale for law professors who aspire to do empirical work but wonder whether they have the background for it. When Chambers

² David L. Chambers, *Making Fathers Pay: The Enforcement of Child Support*, University of Chicago Press, 1979.

³ Sherman, Lawrence W. and Richard A. Berk, "The Specific Deterrent Effects of Arrest for Domestic Assault." 49 *Am. Soc. Rev.* 261 (1964).

⁴ Lenore J. Weitzman, *The Divorce Revolution: The Unexpected Social and Economic Consequences for Women and Children in America*, The Free Press(1985)

began this work, he was a lawyer pure and simple with no special training in empirical investigation. The study's earliest roots were not in Chambers' family law teaching but in his role as an advisor to a law-student group that offered advice on civil matters to inmates at the Detroit House of Corrections. In his role as advisor, Chambers accompanied his students on a tour of the jail. There he saw a number of jail trustees lounging around who did not resemble the typical prisoner stereotype. When he asked what these men were in for he was told by the jailer that these were his "runaway pappies." It turned out that these men were in jail for non-payment of child support.

This was surprising, to say the least, for at the time Chambers asked his question, in the 1970s, the child support system was notorious for its laxity and its failure to collect the money due divorced wives who had become custodial parents. Overall child support collections were estimated to be in the low teens as a percent of what was owed, and the system's reluctance to pursue and punish men who ignored child support enforcement orders was becoming a cause *celebre* of the then developing women's movement. Yet here was a state, or at least a county in a state, that was actually jailing men for failing to meet their support obligations. Chambers decided to take a close empirical look at what was happening.

Recalling early conversations, I think it fair to say that when Chambers began his empirical effort he did not think that jailing men to collect child support did much good, nor did he want it to because he did not like to see mainly poor men who were guilty of only a civil offense serving time in jail. But his early efforts to gather data immediately called this perspective into question. The data available indicated that as compared to other states Michigan was remarkably effective in collecting child support; indeed a highly disproportionate share of the nation's total child support receipts were collected in Michigan, and men would flee Michigan to avoid the state's child support enforcement practices.

This early field work revealed not only that Michigan did well in collecting child support, but also that a particular institution, the Friend of the Court, played an important role in what was happening, for in Michigan, unlike most states, payments did not go from divorced father (or non-custodial parent) to divorced mother (or custodial parent). Rather payments went from the divorced father to the Friend of the Court's office and then to the divorced mother. Thus a state office had immediate knowledge of which

men were behind on their support payments. Friend of the Court offices for some Michigan counties proactively pursued men who fell behind, while others, although they knew who was behind, waited until a mother complained before taking action. Also when they did take action some Friend of the Court offices were aggressive in seeking jail time for delinquent fathers while others worked more with fathers, and, if they used jail at all, used it sparingly.

To shorten what could be a very long story, let me cut to the chase. What Chambers found was that the two variables that distinguished the different county Friend of the Court offices, proactivity and the use of jail, played a crucial role in the ability of counties to collect child support debts. Controlling for other relevant variables, like the father's job situation, counties that pursued delinquent fathers without waiting for an ex-spouse to complain and that moved quickly to jail fathers who persisted in their delinquency were substantially more successful in collecting amounts due than those counties that were proactive but reluctant to jail, reactive though quick to jail once a mother complained, or neither proactive nor willing to quickly jail. Chambers examined the data from both a cross-sectional and time series perspective and he also interviewed some fathers and other key actors, both giving him a richer understanding of problems and processes and allowing him to write a more interesting book.

Chambers also found that many of the jailed fathers had lost jobs, suffered illnesses or for other reasons seemed unable to pay. But even these fathers sometimes paid their child-support debts, apparently because new partners, parents or others were willing to advance money for their release. Chambers concluded his study with a series of policy recommendations designed to reduce, through automatic payroll deductions like those for FICA and income taxes, both child support delinquency and the jailing of fathers. The study provided influential in the then developing debate about how to increase child support collections. Not only did Chambers present his findings to the Congress, but in true testimony to the contributions a lawyer can make through empirical research, he also appeared on the *Today Show* and was featured in *People Magazine*.

Chambers' work is to my mind an exemplar of empirical work by a law professor. He saw an intriguing issue; rather than speculate or argue on principle for his preferences he sought to learn what was actually happening in an area where values conflicted; he designed what was initially a rather simple study; he saw that the complexities of the empirical situation required him to dig more deeply than he originally contemplated; he applied for and

received a substantial grant in a highly competitive process; he brought into the study a more experienced empirical researcher (Terry Adams) to aid with data collection, organization and analysis; he produced a first rate empirical study including results that ran counter to his original hypotheses and value preferences, and explained clearly and argued responsibly for the policy implications that might be drawn from his work.

We also see in this study the importance of understanding mechanism. For example, a study that contrasted Michigan's delinquent father jail rates and child support collection rates with similar rates in the other fifty states would almost certainly have found large differences on both dimensions and, accepting a deterrence theory perspective might easily and convincingly attributed Michigan's child support collection rate to its jailing policies. Yet high jailing without proactivity did little good, and imposed on both men and the state the cost of placing men in jail.

Note also in this example, the importance of the unit of analysis and the potential of getting misleading results when the wrong unit is used. It is only by looking at county level data that one can find differences in proactivity. Since a unit is either proactive or not and counties differed across Michigan with statewide data only this variable would go undefined. Thus the only relevant statewide calculation that would be possible would be the overall state jailing rate since unlike proactivity, county jailing rates could be aggregated to the state level.

Finally, consider the solution Chambers recommended. It is not a solution drawn directly from the data though one might think of it as inspired by it. In advocating not proactivity coupled with jailing but rather automatic payroll deductions, Chambers identified a way of collecting child support that had the virtues of proactivity without the cost of jailing and, which moreover, did not force the obligations of fathers who had no way of meeting them on their family members.

Sherman and Berk: Effects of Arrest on Misdemeanor Spouse Abuse

The study by Larry Sherman and Dick Berk of misdemeanor spouse abuse was, at the time it was done, an exemplary evaluation of law enforcement strategies. It is hard to imagine this study being done by an inexperienced empirical researcher. It required substantial funding, the ability to persuade a police department to cooperate with research in a then almost unheard of way, and sophisticated statistical analyses of the data collected. What Sherman did (Berk, I believe, was brought in later) was to persuade the Minneapolis Police Department to institute a field experiment when called to the scene of alleged spouse (and domestic partner) abuse.

Rather than act on their discretion, as police throughout the nation did, the police were instructed, if the alleged abuse would not support a felony arrest, to open a sealed envelope and, following the instructions inside, to either (1) arrest the man accused of abuse, (2) remove the man from the home but not arrest him or (3) counsel the couple and when things had calmed down to leave. The experiment was implemented more or less as designed (the need for sophisticated statistical analysis resulted from the “less’ part) and yielded seemingly striking results.

The incident that had led to the police being summoned ended with their arrival, no matter what the police did, but whether measured by later arrests or a woman’s response when interviewed by phone six months later, men who had been arrested appeared less likely to engage in repeat abuse than men subject to either of the other two treatments. In part because Sherman made an extraordinary effort to publicize the study’s results (e.g., he arranged to have a T.V. crew ride with the police on one call so that there would be video footage to accompany the announcement of the study’s results when the time came)⁵ and in part because the lesson that arrests works appealed to both conservative law and order types and liberal supporters of the women’s movement, the results of this study reverberated around the country. Numbers of police departments moved to mandatory arrest or other harsher policies for dealing with spouse abuse.

There were, however, problems with the study evident, even in the initial *ASR* article that should have cautioned those who advocated arrest as the presumptive response to misdemeanor spouse abuse.⁶ First, the study had been done in Minneapolis, hardly a typical American city. In particular Minneapolis has, apart from Native Americans, a very small minority population for a city its size. Also Minneapolis soon released the men it arrested, while other jurisdictions might treat arrestees differently. Thus no matter how well done the study, there was reason to fear its results would not generalize to other settings. As importantly, the study said little about the mechanism through which its results occurred. Rather it assumed that its findings reflected a deterrent effect. However, looking at the data, it appeared that the results were largely driven by men who did not return to

⁵ Sherman, Lawrence W. and Ellen C. Cohn, “The Impact of Research on Legal Policy: The Minneapolis Domestic Violence Experiment.” 23 *Law & Soc. Rev.*117 (1989)

⁶ This statement is not made in hindsight. I was editor of the *Law & Society Review* when the study appeared and wrote an editorial both praising the quality of the research and cautioning against relying too heavily on its results. Richard Lempert, *From the Editor* , 18 *Law & Soc. Rev.* 505 (1984)

their domestic partner relationships after being arrested. Perhaps this is a good thing, but perhaps not, and perhaps the choice of whether to take steps that will break up a relationship should be the woman's choice rather than one made through police policy.

The National Institute of Justice which had sponsored the Minneapolis study was a bit taken aback by the speed with which its findings were disseminated and influenced policy. Although proud of this research and pleased with its outcome, NIJ made the wise decision to fund five replications to see if the study's results were robust to variations in treatment and would generalize to other places. Perhaps the best of the replications was done by Sherman and his colleagues in Milwaukee.⁷ Here the results were different from what they had been in Minneapolis. While women of higher socio-economic status, who presumably had partners of higher socio-economic status, did seem to gain protection when their abusers were arrested, women of lower socio-economic status experienced more violence when their abusers had received the arrest treatment. A mechanism suggests itself. Men who had a lot to lose through arrest were deterred from future violence for fear of the sanction. Men who had little to lose or had experience with being arrested were simply angered, and they took their anger out on their partners after release from jail. To put this another way, white, well educated women with employed partners seemed to gain protection from a mandatory arrest policy but this was at the expense of less well educated and black women and women with low status jobs if they worked at all and poorly employed or unemployed partners. Several of the other replications yielded results consistent with Sherman's Milwaukee findings. The policy implications of this body of research are thus quite ambiguous.⁸ Yet the results of the NIJ replications never achieved the popular dissemination of the original Minneapolis study.⁹

Empirical work on this topic has continued. Consider a recent paper by Radha Iyengar.¹⁰ Iyengar uses a different strategy. She uses a quasi-

⁷ See, e.g. Lawrence W. Sherman, Janell D. Schmidt, Dennis P. Rogan, Douglas A. Smith, Patrick R. Gartin, Ellen G. Cohn, J. Collins, Anthony R. Bacich, "The Variable Effects of Arrest on Criminal Careers: The Milwaukee Domestic Violence Experiment," 83 *J. Crim. Law and Crim.* 137 (1992)

⁸ Berk, Richard A., Alec Campbell, Ruth Klap, and Bruce Western, "The Deterrent Effect of Arrest: A Bayesian Analysis of Four Field Experiments," 57 *Am. Soc. Rev.* 698 (1992).

⁹ Arnold Binder, James W. Meeker, "Implications of the Failure to Replicate the Minneapolis Experimental Findings" 58 *Am. Soc. Rev.* 886 (1993)

¹⁰ Radha Iyengar, "Does the Certainty of Arrest Reduce Domestic Violence? Evidence From Mandatory and Recommended Arrest Laws" Working Paper 13186, National Bureau of Economic Research (June 2007).

experimental rather than an experimental approach and looks at domestic partner violence in states with mandatory arrest laws for spouse abuse and states with recommended but not mandatory arrest laws before and after these laws were passed. She finds an apparent increase in intimate partner homicides in mandatory arrest states after they enact mandatory arrest laws but no such increase in recommended arrest states. Her explanation for this is that mandatory arrest laws deter calls to the police who might stop an altercation before it becomes fatal. To test this theory she looks at family homicides that do not involve intimate partners since mandatory arrest laws for family abuse cover such abuse as well but the abused person, often, she speculates, a child will have their cases reported by outside authorities like doctors or teachers. Since homicides in these cases do not increase after the passage of mandatory arrest laws she sees her thesis that spouses are deterred from calling the police by mandatory arrest laws as supported.

This body of research carries numbers of lessons for those interested in using empirical studies to inform social policies. Perhaps the most important is the way political forces can build upon and implement results they find congenial even when much more work is needed to understand a problem. The original Sherman and Berk article carries cautions about generalizing from its findings to policy, but boiler plate cautions in technical articles are unlikely to carry much weight with policy makers, and in this case Sherman's efforts to publicize the research results in the popular press¹¹ more than counteracted any dampening of enthusiasm that might stem from a reference to the need for more research and abstention from hasty action.

Second, even well-designed research has its limits. Well done field experiments are about as good as one gets in policy-related research. But good may be far from perfect, depending on how the research results are applied. One limitation of the Sherman and Berk study is its apparent context dependence. Seemingly similar treatments can have quite different effects in different locations or at different points in time if their implications can depend on population, organization and other local characteristics.¹²

Another limitation of the research from a policy perspective is that despite appearances Sherman and Berk were not testing the deterrent effect of mandatory arrest policies, although the research was interpreted as if they

¹¹ Lawrence Sherman [LSR article]

¹² Totally apart from population differences, cities like Milwaukee and Minneapolis differ on how people are treated after arrest. These may be constant within a jurisdiction, but they vary across jurisdictions and although they are not varied in the experiment they should be considered a part of the experimental treatment.

were. Rather they were testing the experience of three different treatments: arrest, forced separation and cautioning, on men whose abuse they had broken up. No man who was arrested and no spouse of an arrestee had reason to expect an arrest was certain the next time the police were called because of alleged abuse. But when a mandatory arrest law is place, men and women are likely to know it. It is in part this insight that drive's Iyengar's research.

A third point is that wise policy depends on understanding mechanism. In the original Sherman and Berk research it was suggested that the mechanism by which mandatory arrest worked to reduce repeat violence was the deterrence of abuse attributable to the experience of arrest. But there were ambiguities from the start. To begin with there were two measures by which future abuse was measured: records of later calls to the police and reports that the women gave when later interviews. Results across treatments differed by measure. This should have raised if not a red, a yellow caution flag. Perhaps more importantly data that received little attention indicated that arrested men were substantially less likely to return to their intimate partnerships than those who experienced the other treatments. Men may thus have been deterred from maintaining relationships rather than from further abuse. Is the decision to break up relationships where there is some abuse one that the state should make for women, or should women make this decision for themselves, at least when the abuse does not rise to the level of a felony?

Iyengar's study is also ambiguous with respect to mechanism. It could be that, as she suggests, women in mandatory arrest jurisdictions are more reluctant to call the police and so do not use a police summons to extricate themselves from situations likely to become fatally violent.¹³ It may also be the case that men knowing they will be arrested if the police are called are more violent, to the point occasionally of homicide, in attempts to prevent their partners from calling the police. Or the mechanism may be

¹³ Iyengar ignores the fact that the police are often called to the scene of domestic violence by a neighbor disturbed by what is going on or by a child rather than by a spouse. The implications of such calls for her theory and its test are unclear. They could dampen the apparent effects of mandatory arrest laws if her theory is correct, making her results yet more impressive or they could suggest there is a need to look elsewhere for causality and, in particular, that her control for other family (expect to be child) abuse arrests is not as strong as it might appear because the difference in the proportion of calls initiated by the abused person is less than it might seem. What we need is some information about the proportion of spouse abuse calls that come from third persons. The child abuse control is also weak, I might add, because mandatory arrest laws assume the abuser is known. When police arrive while the abuse is ongoing, it is clear abuse has occurred and the identity of the abuser is relatively easy to establish, even if both parties are treated as abusers. With child abuse the abuse is not ongoing and the abuser's identity may not be clear. Moreover, the abuser may be the mother rather than the father, and men may respond differently to arrests for abuse than women.

entirely different. Perhaps women in recommended arrest states or in states with no laws feel more empowered vis-à-vis their men because whether a man gets arrested may depend largely on their preferences. Alternatively the mechanism may have nothing directly to do with the laws impact during an argument. Unless first offense calls to the police are substantially deterred in mandatory arrest states, presumably more men in these states will have been arrested for spouse abuse than men in other states. These arrests may have caused them to be embarrassed in front of colleagues, to have to spend money on bail and lawyers fees or even to lose their jobs. Perhaps it is the strains that these kinds of experiences place on a marriage or the drinking they can lead to that increases homicide rates even if the initial experience of arrest has a deterrent effect.

One difference between Minneapolis and Milwaukee was that in Minneapolis men were soon released after arrest and often suffered few further consequences, but in Milwaukee they were likely to be in jail longer and the collateral effects of the arrest may have been worse. Since states likely to enact mandatory arrest laws may be likely to treat men accused of abuse more harshly in other respects, it could be treatment other than arrest that is at the root of apparent homicide wise. If Iyengar had information on state bail and charging policies she might be able to do more to pinpoint the reasons for the results she reports. In this respect Iyengar's study illustrates a further need for caution. Regression models like those she uses can only control for variables they can measure. If influential variables are not included in the model an explanation may be entirely different from what might appear. A major advantage of the randomized field experiment is that it can control for variables that are not be measured.

Both Sherman and Berk and Iyengar have in their own ways conducted studies that have considerable strengths as social science. The Sherman and Berk study was more than that, for although it was not the first field experiment done to study policing¹⁴ it was path breaking in implementing an experimental design at the individual arrest level. Nevertheless, it is fair to ask what light this research has shed on the question of whether mandatory arrest policies for misdemeanor spouse abuse protect women's interests. I conclude not much. Iyengar may be right in her judgment that on balance mandatory arrest increases rather than decreases intimate partner violence, or as the Milwaukee study suggested mandatory arrest policies may protect some women and harm others, depending on their characteristics and those of their partners. But one can

¹⁴ The first was probably the Kansas City intensive patrol experiment. [Cites]

hardly have a law that calls for the mandatory arrest of people with some demographic characteristics while only warning or separating those in other groups.

This does not mean that empirical research could not productively inform police responses to domestic violence. Indeed, early research might have done that by first revealing how feeble police responses often were and later calling into question certain counseling approaches. But to yield the kind of transformative policy implications that many drew from the Minneapolis experiment, more than an occasional experiment here or quasi-experiment there is needed. Society must invest in long term programmatic research. Something we appear unwilling to do. Moreover, the research is unlikely to get us far if it only asks whether one treatment works better than another. Averages, even if accurately assessed, can hide lots of variance. Rather the aim must be to generate theory and understand mechanisms. Only with such understanding is even good social science likely to get us anywhere in the policy arena.

Weitzman: The Divorce Revolution

Lenore Weitzman's research purported to show that following the passage of California's no fault divorce law divorcing women's income fell relative to their needs by 73% while men's income relative to needs improved by 42% . Weitzman's numbers are striking, and it is hard not to attribute them to changes wrought by the California no fault law since other research in states without no fault showed smaller drops in women's relative well being post divorce (about 33%) and smaller gains for post divorce men (about 15%-20%). Weitzman's figures were consequently picked up by the news media, cited in legislative hearings and figured importantly in debates about adopting and structuring no fault divorce laws.

If the work was influential, it is hard to fault the public for being influenced since Weitzman's book received the American Sociology Association's annual book prize award. Still there were reasons to be skeptical of this work from the outset. First, the most publicized results were based on a relatively small sample (228 respondents divided equally between men and women) of people living only in Los Angeles, with an oversample of wealthier individuals. Also Weitzman was not known as a quantitative researcher, and she could not or would not furnish the data from her study to those who sought to acquire it. More importantly the results were inconsistent with numbers of other studies of the effects of divorce on the economic well-being of men and women, and a 73% drop in well-being

is so large as to almost defy credulity, since 100% would be the maximum fall possible. Eventually, Richard Peterson gained access to the raw records from which Weitzman had worked.¹⁵ He corrected numerous coding errors and reanalyzed the data. He found that women's income to needs ratio fell to 27% following divorce, while men's rose by about 10%, results consistent with or even showing somewhat less gender bias than other similar research. Other work fills out the picture. Pamela Smock, for example, finds immediate post divorce disparities in gender well-being are entirely attributable to women with children, reflecting the fact that women are most often custodial parents and the standard need measure accounts for the number of people in a household.¹⁶ Still others, using different measures, have disputed whether men's economic well being indeed increases after divorce.¹⁷

Peterson's figures and most of the best subsequent work do not change the bottom line: there is a serious problem – especially in regard to the post-divorce economic well being of women with children - that needs to be addressed. But the other research does change our understanding of why the problem exists. Moreover, putting studies together it appears that whether one lives in a fault or no fault state has no implications for how well off marital partners will, on average, be post divorce. Professors, students, legislators and others who examined no fault laws through the prism of Weitzman's research appear to have been seriously misled.

One caution to draw from the Weitzman study is the realization that good empirical research is hard to do, which means one must ask who is doing the research. It is not enough to plug data into a canned statistical package and generate results. To decide on tests and interpret results the logic of hypothesis testing must be understood and there must be good appreciation for what measures mean and data quality issues. A second lesson is that when a quantitative study bears importantly on a policy issue, it is important that the data used be widely available and that different researchers examine the issue. A third concerns the danger of agenda-driven research. Research should be aimed at testing theories not at making points. While I don't know if Weitzman set out to show that no fault

¹⁵ Richard R. Peterson, A Re-Evaluation of the Economic Consequences of Divorce, 61 Am. Soc. Rev. 528 (1996).

¹⁶ Pamela J. Smock, Gender and the Short-Run Economic Consequences of Marital Disruption, 73 Soc. Forces 243 (1994).

¹⁷ Arlee Stroup and Gene E. Pollock, Economic Consequences of Marital Dissolution." 22 J. Divorce and Remarriage 37 (1994)

divorce laws harm women, some have thought this was her aim and even if it wasn't her ready willingness to embrace the result might have blinded her to problems with her data and what it suggested. There is a difference between wielding data like a lawyer, which means providing only that information and those interpretations that advance a client's interest and using it like a social scientist which means using it to advance knowledge by putting favored theories to the hardest possible test and revealing and even calling attention to results that contradict what one would like to claim. Only when empirical research has the latter characteristics should it enter policy debates. But too often it is agenda driven research that seems to speak loudest to policy makers.

In addition to my specific comments on the studies I have reviewed in detail, there are a number of lessons the policy maker or law professor reading empirical work might learn from this brief review. Here are five: (1) Don't rest policy change or analysis on a single study no matter how good it is; (2) Look beyond the researcher's bottom line to other relationships revealed in the data; (3) If results seem too good to be true this may be because they are not true, and (4) Always ask about mechanism; understanding why a situation exists is as important to policy analysis as knowing whether it exists, and (5) no matter how unversed one is in statistics, common sense and close reading can take one a long way.

Smock and Manning: Living Together Unmarried

In addition to using empirical research to inform policy, lawyers can look to empirical work to anticipate legal issues that are here but as yet unacknowledged by the law or are likely to arise in the near future. To continue with family law examples, consider a synthetic article by Pamela Smock and Wendy Manning on cohabitation and its implication for family policy.¹⁸ Smock and Manning are family demographers who have published on a range of subjects. Their article documents the rising age of marriage in the United States over the past 50 years (from a median age of about 20 for women in 1950 and 23 for men to a median of 25 for women and almost 27 for men in 2000) and an equally striking rise in the number of cohabiting households (from less than 500,000 in 1960 to more than 4.5

¹⁸ Pamela Smock and Wendy Manning, Living Together Unmarried in the United States: Demographic Perspectives and Implications for Family Policy, 26 Law & Policy 87 (2004)

million in 2000, with 48% of women in their late 30s in 1995 reporting at least spell of cohabitation.) Moreover, they report, about 1/3 of births in the United States are now occurring outside of marriage with about 1 in 8 children now being born to cohabiting couples, many of them planned. Overall about 40% of cohabiting unions now have children present, just a slight bit under the 45% of married couples who report the presence of children. This situation and these trends portend numerous challenges for the law. Family law scholars are aware of many of these challenges and the law has begun to grapple with some of them, but reliable information on the degree and velocity of these changes has important implications for the kinds of responses that are needed and the urgency with which we work toward them.

Doing Empirical Research

Empirical research relating to most issues outside of law and economics has been dominated to date by social scientists rather than lawyers and the situation is not likely to change. However, as David Chambers' research makes clear lawyers, even those initially untrained in empirical research, can do valuable, high quality empirical work. Moreover, they can see issues ripe for investigation which sociologists, economists and others will not have on their radar screens for investigation. Moreover with the increased interest in empirical studies in the law school world, we are likely to see more legal scholars, particularly younger legal scholars, attempting empirical projects. It is for them I write what follows.

Empirical research on legal issues can take various forms. We are today blessed with numbers of high quality data sets, like those produced by the National Longitudinal Study of Youth, Ad Health, the New Immigrant Survey, the Population Study of Income Dynamics and, of course, the various census surveys. Moreover they are conveniently available on line for exploration and investigation, always accompanied by necessary documentation and sometimes with tools to aid analysis. Complementing these data sets are relatively user friendly statistical packages that allow novices to implement statistical approaches to data that were cutting edge only a few years ago. It will, however, be the rare legal scholar who enjoys a comparative advantage in analyzing these data. Simple relationships like those reported by Smock and Manning will be known from the literature, and more complex investigations are best undertaken by those with statistical knowledge that few law professors possess.

But this does not mean that law professors and their students cannot muck around in these data, looking for relationships that others did not think to examine and posing simple tests of association and causality. Moreover, the world of social science scholarship is today far more interdisciplinary than it was even a decade ago, and there is often a place for lawyers on interdisciplinary teams examining data on legally relevant relationships or, indeed, identifying legally relevant questions that social scientists untrained in the law would miss.

Often, however, the kinds of questions that lawyers seek answers to are not questions that can be addressed entirely, if they may be addressed at all, through existing data sources. Law professors interested in a particular issue will often find that they need to collect their own data if they are to shed empirical light on the problem. There are many ways to collect empirical data to inform law and policy. The right way or ways depends on the problem examined, the questions asked and the hypotheses to be tested. The data in David Chambers study was based largely on the records of Friends of the Court offices and interviews with key informants. Weitzman used survey research and Sherman and Berk instituted a field experiment. Austin Sarat and William Felstiner sat in divorce lawyers offices to observe and record interactions between divorce lawyers and their clients.¹⁹ Legal rulings can be examined empirically through content analysis and the assembly of information on judges, lawyers and parties. Litigation trends may be studied using data collected by state and federal authorities or by jury verdict reporting services. Historical studies may draw on news archives, old diaries and moldy records retrieved from court basements. Laboratory studies can examine the psychodynamics of legal processes. And I am sure I have left some approaches and data sources out.

I have observed during my career that law professors and their students are often intrepid empirical researchers, exploring problems on a shoestring that sophisticated social scientists would not touch without six figure grants. This can in part be done because law students are often enthusiastic and low cost data collectors, coders and research assistants. But before lawyers abandon conventional scholarship and start collecting data, several caveats must be mentioned.

¹⁹ Austin Sarat and William L. F. Felstiner, *Divorce Lawyers and their Clients: Power and Meaning in the Legal Process*, Oxford University Press (1995)

First, empirical research takes time, often far longer than one might have anticipated. In particular projects that involve collecting one's own data, as opposed to using existing data sets, are seldom suitable for the untenured, particularly if human subjects are involved, except when sufficient articles have already been produced to ensure tenure or possibly as part of a competent, experienced research team. IRB approval alone can delay projects by months, if not half a year or more. Moreover, data are never acquired in pristine condition, suitable for immediate analysis. Instead substantial time must be devoted to reconciling inconsistencies and otherwise cleaning the data, with additional time devoted to training coders and coding the data.

Once data are collected, cleaned, coded and otherwise ready for analysis, it may be that they will yield a number of valuable articles, but here the untenured faculty member may run into another problem. Law faculty who don't do empirical research and who are used to putting all their ideas on a topic into one gigantic article, may regard several separate smaller articles that test different hypotheses with the same data set as essentially aspects of the same piece, even though it is common for social scientists to draw many separately published pieces from the same acquired data. Equally frustrating, traditionally oriented senior faculty may fail to recognize the time, thought, skill and scholarship that went into the data collection phase of the project. Thus they may see a junior colleague with four or five shorter articles exploring aspects of data she collected as less productive than a cohort at the same career stage with two lengthy pieces, even though the work needed to collect, organize and analyze the first faculty member's data far exceeds that which can be expected to yield two typical "tenure pieces."

Second while some empirical work can be done on the cheap, often doing it right will cost money – more money than a law school's research fund or other easily accessible sources can cover. This may not always appear to be the case at the start, but as my colleague David Chambers discovered early stages of empirical research may open up more questions than answers or may reveal that definitive answers to the problem that motivated the inquiry cannot be provided without a significantly greater and more expensive effort.

The good news is that there are sources of money for projects that can cost two or even three hundred thousand dollars and up. The bad news is that grants of this size, and even substantially smaller grants, are quite competitive and often hard to get. At the National Science Foundation where I spent four years as Director of the Division of Social and Economic Sciences, funding rates typically ran between 20 and 30% of submitted proposals, and it was common to see proposals funded on only a second or third submission. (Another reason why law professors who must produce articles to meet a tenure time table should be cautious about being the lead or solo investigator on major empirical work, particularly if they lack relevant training and grant-writing experience.) Furthermore, at NSF law professors did not get any special breaks with their submissions, but were held to the same standards with respect to theoretical and methodological sophistication as any other person submitting to a social science program. The same is likely to be true of most funding agencies.

These standards include deep knowledge of relevant theory, methodological and statistical skills adequate to the proposed research, and reason to believe that the proposed research will advance knowledge in a field or discipline. While the potential practical importance of a research project would never count against it and is in fact something of a plus, potential practical value alone would not justify NSF funding. Indeed, panelists reviewing proposals in most if not all the programs I supervised would regard a principle investigator's decision to rest a case for funding largely on the practical value of what might be learned as a sign that the investigator was insufficiently versed in theory and methods or otherwise unconcerned with a project's potential science contributions.

Not all potential funders share NSF's basic science commitment. The National Institute of Justice, which funded the Sherman and Berk study, is, for example, a mission agency. Their mission is supply the science and other resources needed to reduce crime. In the mandatory arrest study, an important driver of funding for them was the likelihood that the proposed field experiment would inform the policing of domestic violence in ways that would discourage repeat violation, but even so the research was presented to them as a test of deterrence theory. Still they were less concerned than NSF would have been with exactly why repeat offending might or might not drop with mandatory arrest than they were with answering the question of whether it would drop, and a randomized field experiment seemed to them the best way of answering the latter question.

But as we saw, inattention to mechanism was the Achilles heel of any otherwise excellent study. Not knowing whether mandatory arrest worked by breaking up cohabitation or through deterrence and not knowing how people in different socioeconomic situations might react to the experience of arrest gave the study a potentially perverse policy relevance that was unfortunately realized in some jurisdictions.

Once a funding source or sources is identified, the task becomes one of preparing and submitting a proposal. Whatever the funder, there will almost always be some person at the funding agency who will be happy to talk about what his or her agency is looking for and how to go about making a proposal. Make an appointment by e-mail to talk to that person, by telephone or even in person if you can manage that and if the person is willing. It never hurts to have the person who will help decide funding feel that an applicant is not a complete stranger or see reflected in a proposal ideas or organization that he or she suggested. As importantly, a conversation can prevent a substantial waste of time, for you may soon learn that despite the language with which proposals have been invited what you would propose funding would either not be eligible for funds or would be unlikely to get funded even if it were not absolutely disqualified.

It is also important to think about the appropriate team for conducting an empirical research project and to check with the funder on its attitudes and concerns regarding team research. While much of the research that NSF funds is single investigator research and while private foundations and other agencies also make single investigator grants, more and more research funding is going to disciplinary partnerships or interdisciplinary teams. Moreover, most law professors seeking funding for empirical work are unlikely to get very far with most funders if they are not working closely with social scientists who can supply the methodological, statistical and other skills that they lack or may appear to lack. In addition experienced social scientists, particularly those with a track record of funding from the agency approached, can be a huge help in tailoring a proposal for an agency. Even if such a person is not part of a team, a law professor principal investigator should try to identify experienced grant getters at his or her university and seek feedback on his or her proposal from them.

Finally we come to the proposal. Exactly what one should write and how it should be organized will vary from funder to funder. The funder's web site and officer in charge can provide relevant advice. Almost always there will be a discussion section which designed to show familiarity with the relevant literature. This section is usually toward the front of the proposal and provides the reviewer with an initial and sometimes unchangeable impression of the applicant's ability to think and write clearly. Don't assign this to a student assistant and be sure the search for relevant prior sources is thorough. I have seen more than one proposal's funding chances go down the drain when a panelist noted that a central article in the field (sometimes by the panelist) was nowhere cited. So if there is a highly regarded article that a reviewer might possibly think relevant to the research, a wise investigator will cite it if only awareness of the work and to explain why it is being disregarded.

Some programs and some panelists like a clear statement of hypotheses in specific, clearly testable form. The need for laying out a proposal in this way should depend on the goals of the research, but some people think good science requires this kind of presentation regardless. Specifying hypothesis does not, however, require having any expectations about what the data will reveal. One can not only specify the null hypothesis which posits that a factor will not matter but he/she can also make clear that there is no basis for deciding in advance whether the null is likely to be rejected. It is also appropriate, particularly in the context of case studies, to indicate that the proposed research is more about hypothesis generation than hypothesis testing, but there is need for care and persuasiveness in making the case.

When I read proposals what most concerns me are the methods for data collection and analysis. One mistake I have frequently encountered is researchers who take so much space reviewing the literature and specifying and justifying hypotheses that they have little room left to discuss and justify their data sources, included variables and analytic methods. No matter how interesting a proposal is to this point, if it is vague or weak on data and methods, I see a decision to fund as buying a pig in a poke. I, and many others I have seen reviewing proposals in panels, will not do this. So researchers must save space in their proposals for an adequately detailed description of how they will acquire the data needed to test their hypotheses and how they will analyze it. This section should make clear the investigators ability to rigorously examine the questions posed in the

proposal with data of adequate quality. This does not mean that weaknesses in the available data or accessible methods should be glossed over. I have seen some proposals turned down for unavoidable weaknesses that a reviewer might have missed had they not been frankly acknowledged, but I have seen many more fall because the failure to perceive a weakness was regarded as a sign that the researcher was not up to the analytic task. Usually one can count on reviewers perceiving the weak points of a proposal. Unless the investigator recognizes them as well, he or she will have no opportunity to persuade the reviewer that they are not as serious as they might seem, that they may be ameliorated to some degree or that to the extent they are inescapable reported results will be properly qualified.

Principle investigators should not lose sleep over the fact that they are inexperienced researchers with no funding history. The admonition to avoid preoccupation with what one cannot change applies here. A track record helps secure funding, but at NSF and many other agencies numerous first time researchers are funded each cycle. Indeed, records are often kept of first time grantees and there can be pressure to ensure that some such applicants are funded. But if one is a new grant applicant, particularly one without a strong history of publications, there is a special burden to prove one's knowledge and capacity. New investigators especially should strive to demonstrate within the body of the proposal that they are well acquainted with the literature, sensitive to the difficulties of the research and knowledgeable about appropriate methods. I still remember one proposal I reviewed when I was a member of the NSF Law and Social Science panel in the late 1970s. No panelist had heard of the principal investigator, as he was just starting his career, but the knowledge and sophistication revealed in the proposal simply bowled us over. Indeed, I think this submission became our number one priority for funding, ahead of many scholars whom we all knew and admired. We were right. The resulting publication is regarded by some now as a classic and the principal investigator has gone on to be a productive scholar of considerable distinction.

Finally, researchers should not expect to get funded on their initial proposal submission. It may happen, but increasingly proposals at NSF and other agencies are funded only after one or more resubmissions. Usually a "declinee" can learn from a proposal's reviews whether resubmission is justified, and the funder's representative – whom the investigator should have talked with several times by now – can usually provide an assessment of the likelihood that a resubmission will be worth additional time and effort.

If one is resubmitting, he/she should not resubmit, essentially the same proposal with mainly cosmetic changes in the hope that a more sympathetic reviewer or panel will now fund what was rejected the first time the first time around. It is not likely. At NSF and elsewhere reviewers and panelists will often include some people new to the proposal but others who have seen it before. When a panelist or reviewer recalls having seen a proposal before and notes that little has changed, it is an almost certain kiss of death. Moreover panelists, program officers and other funders are hugely busy. Think of them. While it may seem to the investigator that it can't hurt to submit an only slightly modified version of a rejected proposal since this requires little work and lightening may strike in the form of weaker competition for funds or a more sympathetic review, the review process is hurt when it is burdened by proposals that have only a slight chance of success or, more likely, no chance at all.

Moreover, to submit an only slightly revised proposal, unless one has been advised that that is all that is needed is simply silly. The reviews tell the investigator what aspects of the proposal most troubled a reviewer or panel, so allow one to address specifically the weaknesses that caused the proposal to flounder. What a golden opportunity. This does not provide a guarantee against frustration, for I have seen proposals revised substantially to meet a reviewer's objections only to receive a new review objecting to the revised approach and suggesting that an approach like the original one should have been used. However, far more common is the reviewer or panelist who, recognizing that his/her suggestions have been taken to heart, is sufficiently impressed (and flattered) to recommend funding.

It is, of course, once a proposal is funded that the hard work begins, and more than once a researcher may regret his or her funded commitments. They impose a timetable, and while short term extensions of a grant are usually easy to arrange so long as additional funding is not sought, there is nonetheless considerable pressure to produce promised products near the time one originally promised to have them. This is also, however, where the sense that what one is doing matters, and the fun of the project begins. Working with empirical data is a bit like solving a mystery. One continually asks questions of the data. More often than not the data cooperate and provide answers, but the answers are seldom final ones. Rather they raise further questions that take one back to the data for new analyses. Fortunately, far more often than not, one does have a product worth sharing with, if not the world, at least one's scholarly community. At this stage

there is nothing like empirical work to give a sense of accomplishment. The writer of a typical law review article, even a very good one, is disseminating ideas with little more than the hope that others will be influenced by his or her analysis and that at some point policy will be impacted. But this rarely happens, even with the work of influential scholars. More commonly a good law review article is celebrated in academic circles for a while and then replaced in people's consciousness by a new hot article. Work that presents the results of sound empirical research, not only communicates ideas, but it adds to the profession's stock of knowledge. Even people with interests very different from those of the investigator or different perspectives on the problem studied can draw on what good empirical research has found. Indeed, collected data may itself be shared, and the sharing may lead to contributions that go far beyond those that the original investigator made. Unlike ideas which, if they hadn't originated in one person may well have originated with others, facts perceived through empirical investigation might have remained unknown had the original investigator not thought to look. In the law as in other disciplines, empirical researchers add to the general stock of relevant knowledge. Although I have posited as a rule of thumb that no one study ever provides a sound basis for policy transformation, a series of consistent studies may well do so. The more deeply legal scholars get involved in empirical research the better the chance that sound, evidence-based laws and polices will emerge.