

Commentary about the Scientific Merit of the Post Release Employment Project (PREP)

Introduction – PREP is an Observational Study

PREP was designed to evaluate the effectiveness of the Bureau of Prisons job training programs (Federal Prison Industries and Vocational Training and Apprenticeship programs). The research design employed is referred to as an observational study. An observational study concerns the effects caused by a treatment or intervention and therefore resembles an experiment (Rosenbaum, 1995). The noted mathematical statistician William G. Cochran introduced observational studies and defined them as empirical investigations of cause and effect relationships applicable when experimental designs are unrealistic or impractical (Cochran, 1953, 1965, 1968; Rosenbaum, 1995). An observational study differs from an experiment in the way subjects are assigned to treatment groups. In an experiment the investigator controls the assignment and ensures that subjects receiving different treatments are comparable. This is generally accomplished by randomized assignments to treatments. In an observational study this control over treatment assignment is absent (that is, it is not achieved via randomized assignment) for any number of reasons. The reasons may be political, legal, ethical, or perhaps practical in nature. Given the choice between a randomized experiment and an observational study one would prefer an experiment. Although inferior to an *equivalent* experiment, an observational study may be superior to a marginally relevant experiment or to a broken (poorly executed or haphazard) experimental design (Cochran and Rubin, 1973; Barnard, 1998).

An evaluation of the Bureau's job training programs is an instance where a randomized experimental design would have been both impractical and unethical, and perhaps may have created legal challenges as well. Random allocation to these programs would have been impractical because in the history of Federal Prison Industries the availability of an industries or vocational job has been based on first come first serve. Furthermore, this evaluation was initiated long after the programs were put into place in virtually every federal facility in the nation, and consequently a full workforce and cue of interested future employees were well

established. It would have been impossible, or at least impractical, to orchestrate and simultaneously implement a randomized selection in every facility, with no way of insuring that the randomization had indeed been implemented as dictated by the research design. Taking people out of their job or placing others into a job based on random allocation may well have had other unintended consequences for the operations of the facilities. Random allocation would also have posed ethical problems, ones that would likely have been remedied by inmates contacting either the courts or their congressional representative, which would have in the end rendered the random assignment moot.

When random assignment is not feasible, program effects can be assessed by observational studies using matched sampling techniques to construct a quasi-experimental control group. Matched sampling refers to selecting comparison subjects (non-program participants) one at a time so as to be similar to individuals who have been treated (program participants) in terms of personal characteristics measured prior to the treatment (Rosenbaum, 1995; Rubin, 1979). PREP used a matched sampling method refined by Rubin (1973, 1979) and Rosenbaum and Rubin(1983, 1984, 1985).

At the time that PREP was being conducted the use of matched sample designs in evaluations of programs with human subjects were less common than in more recent years. The growth in the number of observational studies using matched sampling seems to indicate an increased confidence in the scientific merit of these methods. To our knowledge, at that time, there were no other applications of this method to evaluations of adult correctional programs. Other evaluations of adult correctional programs of that era used flawed experimental designs that did not control for confounding or corruptions in their intended design due to, for example, self-selection bias (resulting from subjects who themselves decided whether they wanted to be in the treatment group or not by virtue of their initial or subsequent decision regarding program participation). Or these evaluations employed non-experimental designs that did not utilize matched sampling methods to construct a comparison group (for example, constructing a comparison group from a random or convenient sample of the general population of inmates).

It has been demonstrated that matched sampling often creates treatment and control groups that are comparable in terms of the variables used in the matching, though the groups may still differ in other ways (Rosenbaum, 1995; Rubin, 1973; Rosenbaum and Rubin, 1983,1984). The degree of comparability between the groups created by the matched sampling method is determined by the relevance of the array of measures or characteristics of individuals used in the matched sampling process. The relevance of these measures is established by the potential for these characteristics to determine the outcome being studied, irrespective of treatment involvement. If such characteristics can be identified and empirically measured, then matching on these characteristics will balance the groups with respect to the prevalence of these characteristics and consequently reduce the bias that might accrue from a non-randomized experimental design that would potentially leave the groups unbalanced with respect to this set of characteristics.

The adequacy of observational designs, and the use of matched sampling strategies for constructing relevant comparison groups, has been reviewed and evaluated by a large number of scientists and scholars (for example, Cochran and Rubin 1973; Rubin 1979, Rubin and Rosenbaum 1983, 1984; Heckman 1979; Deja and Wabah 1999). The remainder of this document surveys the adult corrections job training literature that has referenced PREP. Some of these studies have cited the PREP findings. In other cases researchers have evaluated the quality of the PREP research design and its validity.

Comparisons of PREP with PIRP

A number of studies have compared the PREP results to those obtained from a study of the effects of prison industries experience in New York state prisons. That study, the Prison Industry Research Project (PIRP) was reported by Maguire, Flanagan and Thornberry (1988). In that publication Maguire et al. report that they could not find a statistically significant difference in recidivism rates between those who worked in prison industries and those who did not work in prison industries.

We are aware of three reports that comment on the inconsistency in the findings of the PIRP and

the PREP. In a report (Study # 206) prepared for The National Center for Policy Analysis entitled “Why the Private Sector Needs to Be Involved” (Reynolds, 2001), the author, Morgan O. Reynolds, Professor of Economics at Texas A&M University, cites the encouraging findings of the PREP study, but in a footnote provides a caveat that other studies, like PIRP have not been able to demonstrate similar affects on recidivism. (Professor Reynolds also cites the PREP results without the footnoted caveat about the inability of the PIRP study to obtain findings similar to those of PREP (Reynolds, 1997).) Professor Steven D. Levitt of the Department of Economics, University of Chicago also commented on the lack of consistency in the reported effects of prison industries jobs when the outcome is recidivism, and expressed concern about the possible influence of self-selection on the results reported by PREP and similar studies of inmates in Florida and Ohio. The third study entitled “Federal Prison Industries: Occupational Training or Slave Labor?” by Gary D. “DCDave” Martin, Ph. D. (1997, 1998) also draws a comparison between PIRP and PREP, and the author prefers the PIRP results to those observed via PREP.

In the concluding remarks of Professor Levitt’s report entitled “Preliminary Opinion on the Economics of Inmate Labor Participation” which was prepared for The National Symposium: The Economics of Inmate Labor Force Participation, held at George Washington University on May 21, 1999, Professor Levitt wonders whether the differences that were observed in the PREP are real or whether they might be attributable to a lack of comparability between the work participation group and the non-work group. As evidence for his conjecture Professor Levitt cites the Maguire et al. finding. Specifically, that once they had statistically controlled for a variety of inmate characteristics, such as the number of prior felony arrests and so forth, the recidivism rates between study and comparison group members were statistically indistinguishable. As Professor Levitt did not comment on the matched sampling methods that we used to control for self-selection bias perhaps he was unaware of that facet of the PREP research design. Prior research has consistently demonstrated that matched sampling methods (like the one we used in PREP) are superior to statistical adjustments alone, when these adjustments are applied to a none matched sample of subjects drawn from the same population that contained the program group members (this was the method used in PIRP), (see, for

example, Cochran and Rubin, 1973; Rubin, 1979, or Rosenbaum, 1995). The issue that Professor Levitt raises is certainly the question for which we need a convincing answer if we are to demonstrate the validity of the PREP findings. Indeed, we have attempted to provide a compelling answer in the documents we have written describing that research project. Other scholars who have reviewed and commented on the methods we employed to control for this type of bias have regarded the methods as adequate or even exemplary.

It is likely that the differences in program effects observed in PIRP and PREP could be explained by one or more of at least four factors: 1) differences in the populations of New York State and Federal inmates, 2) differences in the types or qualities of the programs that were evaluated in each system, 3) differences in the job opportunities, and other community resources and social climates experienced by the former New York state inmates as compared to former Federal inmates, or 4) differences in the research methods used to control for self-selection bias due to the non-randomized observational design employed by both studies.

In a November 10, 1991 letter to us from Professor Timothy J. Flanagan, the principle investigator of PIRP (Flanagan, 1991), he offers the following observations on his PIRP and the Bureau's PREP studies.

“Our study, the Prison Industry Research Project, or PIRP, was designed to examine the same basic questions that the PREP project is exploring – whether participation in prison work programs has discernable “effects” on inmate behavior.

[A] critical difference between our study in New York and the PREP study was that the New York research was cross-sectional, retrospective analysis, while PREP was designed to be longitudinal and prospective. For the purposes of your work, your design was far superior to the approach taken in the PIRP study, or any existing study of the impact of prison work on offender behavior.

You and your colleagues correctly identified the problem of self-selection and the potential selection bias as the most significant threat to the validity of your findings. We of course had similar concerns in our New York research. However, because we used the

retrospective design, we used a different method to attack the problem.

The approach that you and your colleagues employed in the PREP project to handle the problem of selection bias was much more sophisticated and powerful than the methods used in previous research. I have examined the propensity scoring procedure in some detail, and I'm convinced that it represents an effective approach to the challenge of selection bias in studies of this type. Most would agree that a pure "experimental design" incorporating random assignment to treatment and control groups would be preferable from a scientific standpoint, but would acknowledge that such designs are all but impossible to carry out in the criminal justice system. There are an overwhelming number of administrative, legal, and ethical objections to "experimenting" with prisoners.

[T]he Post Release Employment Project fits squarely within the modern tradition of correctional outcomes research. In contrast to the universal maxim that "nothing works" in correctional programming, sophisticated studies such as PREP show that some programs can be effective for some offenders under some circumstances. The challenge to correctional policy makers and administrators is to use this research wisely to inform program development and implementation. The Post Release Employment Project is producing valid and reliable information on which to base these decisions."

Other Scientific Citations of PREP

In another letter dated October 14, 1992, to us from the distinguished criminologist Daniel Glaser, he states to us "Your office is doing a good job, and I particularly like the PREP study...Keep up the good work! I'll be glad to see any other reports your office completes of general interest to the public, and to comment on drafts and proposals if you wish."

Several other highly esteemed scholars also presented papers at the George Washington University Symposium on The Economics of Inmate Labor Force Participation, held in May of 1999 (Professor Richard B. Freeman, Chair of Economics, Harvard University; Professors Jeffery Kling and Alan B. Krueger, both of the Economics and Public Affairs Department of the

Woodrow Wilson School of Public and International Affairs, Princeton University; and Ray Marshall, Professor Emeritus of Economics and Public Affairs at the Lyndon B. Johnson School of Public Affairs of the University of Texas at Austin). Although these scholars did not comment on the matched sampling methods we employed they did cite the PREP findings. It seems reasonable to infer that these scholars were comfortable with the methods and the validity of the findings or they would not have been inclined to cite the results without reservations or caveat.

In a subsequent paper entitled “The Labor Market Consequences of ‘Mass’ Incarceration” (Kling et al., 2000) presented at the Urban Institute Re-Entry Roundtable which was held on October 12-13, 2000, professors Kling, Weiman and Western again comment on the PREP study. In this paper Kling et al. “discuss selected examples of the methodologically strongest studies”. Their selection process identified four studies which assess the affect of correctional job training programs on post release employment and recidivism. One of the four studies is the PREP. They state that “More recent evidence has been somewhat more encouraging. Two of the better studies [Project Rio in Texas and PREP], while not using random assignment, have attempted to construct comparison groups of non-participants against which to gauge the effects of the interventions...A study of federal prisoners who participated in academic, vocational and work experience programming used a much more extensive set of variables and more sophisticated statistical matching techniques”.

In a recent publication Uggen and Staff (2001) review the research on the relationship between work and crime in order to ascertain whether employment is a “turning point” in the lives of criminal offenders. The authors “highlight the major employment and training initiatives for ex-offenders in the past 30 years using randomized assignment or strong statistical correction techniques to account for self-selection.” Their list of major initiatives is composed of 10 studies, 6 with randomized assignment to programs, 3 with randomized assignment and matched comparison, and one, the PREP study, with non-random assignment with statistical corrections. With regard to PREP the authors say “Even though inmates self-selected into the program,

Saylor and Gaes (1997) used a sophisticated propensity score methodology to partially address the selection process and reduce bias in estimates of program effects.”

Another evaluation of the PREP methods has been provided by Royce Crocker, a member of the Library of Congress Congressional Research Service Office. On Thursday August 6, 1998, the day after a Congressional hearing on Federal Prison Industries’ mandatory source provision, we received a call from Mr. Crocker. He had been directed to evaluate the research methods employed in the Post Release Employment Project (PREP) and wanted to ensure that he had all of the relevant documents. He did indeed have a copy of every report we had produced, and it was apparent that he had already closely read each of them, as he had considerable knowledge of their content. He impressed us as someone who had a reasonable command of scientific principles and of statistical methods (he had already obtained copies of the Journal of the American Statistical Association articles cited in the PREP documents and seemed knowledgeable about their content as well). He seemed satisfied with the answers to his questions and with the statistical summaries that were sent to him. He expressed that he was impressed by the quality and sophistication of the study and its methods, which he said were far superior to any other government research that he had reviewed (Crocker, August 6, 1998).

In a subsequent phone conversation between Royce Crocker at the Congressional Research Service Office and Chris Erlewine, then Deputy Assistant Director for the Bureau of Prisons’ IPPA Division, which has responsibility for congressional affairs, Mr. Crocker indicated that a Congressional office had asked CRS to review the research underlying the PREP study report. He explained that his role in this situation was more of an interpreter rather than critic, but also volunteered that he was very impressed with the sophistication of the research. In fact, he was surprised to find that he needed to bone up on a number of statistical areas before he was able to thoroughly review the study. Mr. Crocker indicated that he had been reviewing government studies for 20 years, and this was one of the best he had ever seen. He said he was most impressed by the fact that BOP was able to get the various different parts of the organization to cooperate so well in carrying out the study (Crocker, August 21, 1998).

Meta-Analyses of Job and Job Training Program Evaluations

In recent years there have been an increasing number of studies of the program evaluations themselves. These studies, referred to as meta-analyses, or research syntheses, are intended to provide an understanding of how program evaluations arrive at their conclusions. That is, assess the consistency with which evaluation outcomes are associated with the methods employed in program evaluations. The term meta-analysis refers to “ the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings” (Glass, 1976; also, see Cooper and Hedges,1994). And, as Olkin and Hedges (1985) add, “Because meta-analysis usually relies on “data” in the form of summary statistics derived from the primary analyses of studies, it is truly an analysis of the results of statistical analyses”. All of the meta-analyses of evaluations of adult prison jobs or job training programs that we are aware of have included the PREP among the studies they have analyzed.

Bushway and Reuter have a chapter entitled “Labor markets and crime risk factors” in the Sherman et al. (1997) report to The United States Congress. This report was funded by and prepared for the National Institute of Justice, and is entitled “Preventing crime: what works, what doesn’t, what’s promising”. The chapter presents a very comprehensive meta-analysis of evaluations that have been conducted on in prison and post-release job training programs. All of the meta-analyses contained in this report to congress employ a “methodological rigor” rating based on a scale adapted from one used in a national study of the effectiveness of substance abuse prevention efforts that was conducted by the Center for Substance Abuse Prevention. The scientific methods scale rates seven dimensions of the methods used in each study. Based on this scale, each eligible study examined for this report was given a score of 1 to 5, where 5 represents studies with the strongest scientific evidence. Both the PREP and PIRP studies were assigned a score of 3. The authors state that “In order to reach level 3, a study had to employ some kind of control or comparison group to test and refute the rival theory that crime would have had the same trend without the crime prevention program; it also had to attempt to control for obvious differences between the groups, and attend to quality of measurement and to attrition issues”. In the Bushway and Reuter analysis no other reference is made to the PIRP study, however, they

state that the PREP study is “one of the better studies of prison industry programs”. Although the authors express disappointment that the program participation is so broad that it is hard to determine which program or program element led to the decline in re-arrests they nevertheless state that “this seems to be clear evidence that vocational education in federal prisons helps to reduce crime. This is an important positive contribution.”

Wilson et al. (1999) and Wilson et al. (2000) report another meta-analysis of 33 studies involving evaluations of educational, vocational, or work programs for convicted adults in prisons or jails, or other corrections-based program, such as probation. PIRP and PREP were among the studies included in both meta-analyses. In the first of the two published reports the investigators find the results of their analysis promising (participants recidivated at a lower rate than did non-participants) but found that the methods employed in the evaluations of these programs were typically weak. Consequently they caution that there is insufficient evidence to draw strong conclusions regarding the effects of correctional work programs on future offending rates for program participants. Nevertheless, they go on to say

“there were, however, two exceptions. Lattimore et al. conducted a randomized evaluation of an integrated vocational training and reentry program with close to 300 participants and found that participants in the integrated program had a lower rate of re-arrest than did control group members. Saylor and Gaes, using a high-quality propensity score-based quasi-experimental design, evaluated a vocational training and apprenticeship program, also finding positive program effects...Although these two evaluations provide strong evidence that vocational programming can reduce post-release offending, both represent better-integrated and more intensive programs than are typically found throughout the criminal justice system. Thus the generalizability of this finding to typical vocational programs is unwarranted ”. Furthermore, the authors state that *“Although not a randomized design, the study by Saylor and Gaes used the propensity score method to adjust for selection bias and as such, it represents a high-quality quasi-experimental design. Although the propensity score method is unlikely to account for all of the important differences between those offenders participating in work*

programs and those not participating, the careful attention of this study to the problem of selection bias reduces the likelihood that the finding is an artifact of the participation process rather than an effect of the program.” (Wilson et al., 1999).

In the subsequent published report the authors state “The typical comparison group in these studies was a sample of naturally occurring nonparticipants... The only control variables were generally restricted to gender, race, and age. Only one non-randomized study (Saylor and Gaes 1996) used what we considered a strong quasi-experimental design that statistically controlled for important sources of selection bias between participants and nonparticipants, such as prior criminal history, in the analysis of recidivism” (Wilson et al., 2000).

Another very ambitious and comprehensive meta-analysis was recently conducted by the Washington State Institute for Public Policy, an agency created by the Washington State legislature to carry out practical research on issues of importance to Washington State (Aos, 2001). The authors systematically analyzed evaluations of correctional program interventions for juveniles and adults produced in North America over the last 25 years, in order to determine whether the program benefits are likely to outweigh program costs. In the process the Institute systematically reviewed over 400 research studies. As with the University of Maryland meta-analysis (Sherman et al., 1997) this meta-analysis also incorporates a measure of the quality of the research designs employed by the evaluations they included in their meta-analysis. Indeed they used a modification of the scale that was used by the University of Maryland study. Consequently this scale also had a range of 1 to 5, with a score of 5 indicating studies of the highest quality, reflecting evaluations in which the most confidence can be placed. A rating of 4 was assigned to PREP, while a rating of 3 was assigned to PIRP. The institute provides the following definitions for numerical ratings of 3 and 4.

“A level “3” indicates an evaluation where the program and comparison groups were matched for pre-existing differences in key variables. There must be evidence presented in the evaluation that indicates few, if any, significant differences in these variables. Alternatively, if an evaluation employs statistical techniques (e.g., logistic regression) to control for pre-existing differences, and if the analysis is successfully completed, then a

study with some differences in matched pre-existing variables can qualify as a level 3 study.”

“A level “4” is assigned to a study that employs a quasi-experimental research design with a program and matched comparison group, controlling with statistical methods for self-selection bias that might otherwise influence outcomes. These methods may include an instrumental variables or Heckman approach to modeling self-selection. A level 4 study may also be used to represent an experimental random assignment design that had problems in implementation, perhaps with significant attrition rates.”

The institute found only two well designed evaluations of in-prison vocational education and only three well designed evaluations of correctional industries that met their minimum research quality standards for inclusion in their study. PREP was one of the two vocational studies and one of the three correctional industries studies. The institute’s conclusion was that there is a sizable return to the taxpayer for the investment in these vocational and industries programs. (The report provides very specific estimates of the dollars per participant that are returned to society due to an inmate’s involvement in a program .)

Conclusion

Reports of the various stages and facets of the PREP study have been available to the public, the academic community, members of the U.S. Congress, state Congresses, and state departments of corrections for over a decade. The results have been reviewed and commented on by individuals in each of these communities, and they have been overwhelmingly favorable.

April 19, 2002

William G. Saylor, Deputy Chief

Gerald G. Gaes, Chief

Office of Research and Evaluation

Federal Bureau of Prisons

References

- Aos, Steve, Polly Phipps, Robert Barnoski, and Roxanne Lieb. 2001. *The Comparative Costs and Benefits of Programs to Reduce Crime: Ver. 4.0*. Washington State Institute for Public Policy. (Doc # 01-05-1201). <http://www.wa.gov/wsipp>.
- Barnard, John, Jiangtao Du, Jennifer L. Hill, and Donald B. Rubin. 1998. "A Broader Template for Analyzing Broken Randomized Experiments". In *Sociological Methods & Research* (ed. C. Winship). Thousand Oaks, CA: Sage Publications, Inc.
- Bushway, Shawn, and Peter Reuter. 1997. "Labor Markets and Crime Risk Factors". In Lawrence W. Sherman et. al., *Preventing Crime: What Works, What Doesn't, What's Promising*. <http://www.ncjrs.org/works>
- Cochran, W. 1983. *Planning and Analysis of Observational Studies*. New York: Wiley.
- Cochran, W. G. 1965. "The planning of observational studies of human populations (with Discussion)". In *Journal of the Royal Statistical Society, Series A*, 128:134-155. Reprinted in *Readings in Economic Statistics and Econometrics* (ed. A. Zellner, 1968). Boston: Little Brown, pp. 11-36.
- . 1968. "The effectiveness of adjustment by subclassification in removing bias in observational studies". *Biometrics* 24:205-213.
- Cochran, W. G., and D. B. Rubin. 1973. "Controlling bias in observational studies: A review". *Sankhya, Series A*, 35:417-446.
- Crocker, Royce. (personal communication to Chris Erlewine, Aug. 21, 1998).
- Crocker, Royce. (personal communication to William Saylor, Aug. 6, 1998).
- Flanagan, Timothy J. (personal communication to Gerald Gaes, November 10, 1991).
- Freeman, Richard. 1999. "Making the Most from Prison Labor". *The Economics of Inmate Labor Force Participation*. Paper presented at the National Symposium, George Washington University, Washington, DC.
- Glaser, Daniel. (personal communication to William Saylor, October 14, 1992).
- Glass, Gene V., Barry McGaw, and Mary Lee Smith. 1981. *Meta-Analysis in Social Research*: Sage Publications, Inc.

- Hedges, Larry V., and Ingram Olkin. 1985. *Statistical Methods for Meta-Analysis*. San Diego, CA: Academic Press, Inc.
- Kling, Jeffrey R., and Alan B. Krueger. 1999. "Costs, Benefits and Distributional Consequences of Inmate Labor". *The Economics of Inmate Labor Force Participation*. Paper presented at the National Symposium, George Washington University, Washington, DC.
- Kling, Jeffrey, David F. Weiman, and Bruce Western. 2000. *The Labor Market Consequences of 'Mass' Incarceration*.
- Levitt, Steven D. 1999. "Preliminary Opinion on the Economics of Inmate Labor Participation". *The Economics of Inmate Labor Force Participation*. Paper presented at the National Symposium, George Washington University, Washington, DC.
- Marshall, Ray. 1999. "Preliminary Opinion on the Economics of Inmate Labor Participation". *The Economics of Inmate Labor Force Participation*, Paper presented at the National Symposium, George Washington University, Washington, DC.
- Martin, David. (1997, August 9). "UNICOR V". *DCDave's Column*.
<http://www.thebird.org/host/dcdave/article1/080997d.html>
- Martin, Gary D. 1998. "Federal Prison Industries: Occupational Training or Slave Labor". Subcommittee on Oversight and Investigation Committee on Education and the Workforce, U.S. House of Representatives. (Testimony).
<http://www.southernautomotive.com/fpi/martin.htm>
- Reynolds, Morgan O. 1997. *The Economic Impact of Prison Labor: National Center for Policy Analysis*.
- . 2001. *Factories Behind Bars: National Center for Policy Analysis*.
- Rosenbaum, Paul R. 1995. *Observational Studies*. Springer-Verlag New York, Inc.
- Rosenbaum, P., and D. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects". *Biometrics* 70:41-55.
- . 1984. "Reducing bias in observational studies using subclassification on the propensity score". *Journal of the American Statistical Association* 79:516-524.

- . 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score". *American Statistician* 39:33-38.
- Rubin, D. B. 1973. "The use of matched sampling and regression adjustment to remove bias in observational studies". *Biometrics* 29:185-203.
- . 1979. "Using multivariate matched sampling and regression adjustment to control bias in observational studies". *Journal of the American Statistical Association* 74:318-328.
- Sherman, Lawrence W. "Thinking About Crime Prevention". In *Preventing Crime: What Works, What Doesn't, What's Promising*. (ed. L. W. Sherman et. al.) <http://www.ncjrs.org/works/chapter2.htm>
- Sherman, Lawrence W., Denise Gottfredson, Doris MacKenzie, John Eck, Peter Reuter, and Shawn Bushway. *Preventing Crime: What Works, What Doesn't, What Promising*. College Park, MD: Department of Criminology and Criminal Justice, University of Maryland. <http://www.ncjrs.org/works>
- Uggen, Christopher, and Jeremy Staff. 2001. "Work as a Turning Point for Criminal Offenders". *Corrections Management Quarterly* 5 (4):1-16.
- Wilson, David B., Catherine A. Gallagher, Mark B. Coggshall, and Doris L. MacKenzie. 1999. "A Quantitative Review and Description of Corrections-Based Education, Vocation, and Work Programs". *Corrections Management Quarterly* 3 (4):8-18.
- Wilson, David B., Catherine A. Gallagher, and Doris L. MacKenzie. 2000. "A Meta-Analysis of Corrections-Based Education, Vocation, and Work Programs for Adult Offenders". *Journal of Research in Crime and Delinquency* 37 (4):347-368.
- Various. 1994. *The Handbook of Research Synthesis*. Edited by H. Cooper and L. V. Hedges. New York, NY: Russell Sage Foundation.
- The Handbook of Research Synthesis*. 1994. Edited by H. Cooper and L. V. Hedges. New York, NY: Russell Sage Foundation.