

7/11/88

TOWARD MORE VALID EVALUATIONS
OF
TRAINING PROGRAMS SERVING THE DISADVANTAGED

John H. Bishop
Cornell University

Working Paper # 87-07

Center for Advanced Human Resource Studies
New York State School of Industrial and Labor Relations
Cornell University
Ithaca, NY 14851-0952
607/255/2742

This paper has been stimulated by my membership on the Advisory Panel for the National JTPA Evaluation. The research reported here was supported by funds from Contract No. 99-4-576-77-091-01 of the Employment and Training Administration, U.S. Department of Labor. I would like to thank Kevin Hollenbeck for comments on the TJTC empirical work, John Boudreau and John Hunter for their assistance in tracking down studies of productivity variation of workers who are paid the same wage and Robert Hutchens and John Gary for comments on an early version of this paper. The opinions and conclusions expressed herein are solely those of the author and should not be construed as representing the opinions or policies of any agency of the United States Government.

This paper has not undergone formal review or approval of the faculty of the ILR School. It is intended to make the results of Center research, conferences, and projects available to others interested in human resource management in preliminary form to encourage discussion and suggestions.

ABSTRACT

The paper challenges the widespread assumption that the wage effects of federal training programs are reliable and unbiased estimates of productivity effects and social benefits. Evidence is presented that the reputations of government training programs are unreliable and that employers stigmatize those eligible for TJTC and CETA OJT contracts. Graduates of classroom training programs which are known to be funded by JTPA are likely to be similarly stigmatized. TJTC eligibles are seriously underpaid by employers and JTPA graduates may experience a similar fate. Consequently, the true effects of JTPA on the productivity of disadvantaged workers may be considerably larger than its effects on wages. Methods of obtaining estimates of productivity effects are described.

8/18/88

TOWARD MORE VALID EVALUATIONS OF TRAINING PROGRAMS

SERVING THE DISADVANTAGED

By what criteria should the effectiveness of training programs be evaluated? Among policy analysts there appears to be a consensus that these programs should be judged by their ability to increase the earnings of trainees.¹ If educators or employers were asked this question, however, they would probably emphasize instead the training program's "ability to raise the productivity of the trainees." Upon hearing this response policy analysts might say "There is no disagreement. Conceptually productivity is the proper social criterion. Earnings are simply the way of making productivity operational. Estimates of wage effects (produced by a randomized experiment) are unbiased estimates of productivity effects."² But are the wage effects of training initiatives serving the disadvantaged reliable estimators of their productivity effects? Are they unbiased estimates of their productivity effects? I conclude that the answer is no.

In benefit/cost analysis, it is conventional to assume that individuals are paid wages (W_i) equal to their individual marginal revenue products (MRP_i). Yet this assumption is not an implication of modern economic theory. Implicit contracts and signaling theories, for example, provide plausible models of labor markets in which $W_i \neq MRP_i$.

A simple way to test the $W_i = MRP_i$ assumption is to collect data on the productivity of workers at a particular firm who do the same job and are paid the same hourly wage. If worker productivity varies substantially in most samples, $W_i = MRP_i$ must be rejected. The industrial psychology literature contains numerous studies of the variability of output across workers. Recent

reviews of this literature have found that the ratio of the standard deviation of yearly output to mean output, the coefficient of variation, averages about 14.4 percent for semi-skilled blue collar workers, about 16 percent for routine clerical jobs, 28 percent in clerical jobs requiring some decision making and about 35 percent for sales clerks.³ These studies have found that the jobs paid on an hourly or weekly basis typically have larger coefficients of variation than jobs paid on a piece rate basis⁴. With coefficients of variation of this magnitude, it is quite clear that $W_1 = MRP_1$ cannot be true for most jobs.

Are the Reputations of Government Training Programs Reliable?

The assumption that individuals are paid their marginal revenue product has been shown to be invalid. There is, however, another assumption -- wages equal expected marginal revenue product -- which under some circumstances can justify using wage effects of training as a measure of productivity effects. Under this scenario employers do not know the productivity of individual job applicants, but productivity becomes somewhat visible after a worker has been hired. Wages only partially adjust to reflect the perceived productivity differences between workers at the firm. Empirical evidence suggests that after one year, the elasticity of relative wages to relative productivity is .22 in very small firms and zero at establishments with more than 400 employees.⁵ Wages do not adjust completely because workers are risk averse⁶, because measures of job performance are unreliable⁷, because productivity differentials are specific to the firm or only visible to the firm⁸, because large differentials might reduce worker cooperation⁹ and because performance rewards come in the form of small permanent increments to one's wage and not as an immediate bonus. Because follow-ups of the graduates of

training programs seldom extend for more than a year or two, the long delays before persistent performance differentials generate significant wage differentials make evaluations of these programs particularly problematic.

According to this scenario, information on worker productivity does influence the firm's future hiring decisions, however. If high school graduates do a better job on average, then the employer gives hiring preference to high school graduates. If a training program refers particularly capable individuals, then the firm returns to that program for referrals. If referrals from a training program are unsatisfactory, then it does not go back. This scenario envisions employers sharing what they learn about particular training programs with each other. Through this mechanism, training programs develop local reputations that in turn influence the wage offers their graduates receive. The graduates of programs with strong reputations are able to obtain higher wage jobs. The graduates of programs with weak reputations have greater difficulty finding work and typically end up in lower wage jobs.

The reputational mechanism should work reasonably well for visible nonspecific educational achievements such as high school graduation or clerical training in high school. The mechanism might also be expected to work reasonably well when stable education and training programs cater to a limited group of employers (eg., auto mechanics or beautician training) so that the reputations that would develop would be reasonably reliable and competition would force wage rates to reflect these reputations.

The reputational mechanism does not work well in a number of situations, however. First, local labor markets for specific occupations are sometimes monopsonistic or oligopsonistic (eg., training in health care occupations in a city with only one or two hospitals). Customizing training for particular

employers is becoming more popular and this often limits the number of firms at which the training is useful. Under these circumstances training programs produce skills that are to some degree firm specific and one would expect wage effects of training to be smaller than productivity effects.

A second problem is that training programs generally disperse their graduates widely. No single employer has enough contact with the graduates of a specific program to draw a valid conclusion regarding the quality of the training. Because there is no low cost mechanism for pooling the experience of many different employers, reputations are created and evolve very slowly.

A third problem is that many training programs are not stable. This is especially true of the federally funded training initiatives such as Comprehensive Employment and Training Act (CETA), Welfare Work Incentive (WIN), Supported Work and Job Training Partnership Act (JTPA) which have been the subject of the majority of benefit-cost studies. These programs have had very high staff turnover and have undergone frequent changes in design and administration. The reputation developed under one administrative regime may not be valid under another regime.

A fourth problem is that reputations can be manipulated by public relations strategies and are often influenced by idiosyncratic events that receive a lot of publicity. The more remarkable (ie., idiosyncratic) a story is, the more likely it is to be picked up by the national press. The good reputation of the Job Corps was enhanced by the pictures of George Forman dancing around the boxing ring waving a small American flag after winning the Olympic Gold Medal. The reputation of CETA was severely damaged by the

national coverage given to a specific local work experience program that was paying slum youth to play basketball.

Employers of unskilled workers devote few resources to hiring. For retail, service and blue collar jobs, they review on average only 8 applications, interview only 4 applicants and invest a total of 7 hours in the selection process.¹⁰ The interview, the primary method for making selections, has been found to have low reliability and low validity.¹¹ Hiring mistakes are frequent. It seems somewhat farfetched to assume that employers who generally invest so little in their hiring selections have been motivated to overcome the high costs of collecting and disseminating information about local federally sponsored training programs. Increasing the number of trainees studied does not solve this reliability problem, for employer beliefs about the likely productivity of JTPA trainees are not arrived at independently. The job and wage offers received by JTPA trainees are influenced by the general reputation of CETA and JTPA and there are reasons for doubting the reliability of these general reputations.

Not only are wage effects unreliable measures of the productivity effects of federal training programs, they are probably biased as well. A systematic bias results from the fact that participating in a program often informs employers that the trainee is a member of a stigmatized target group and this lowers employer perceptions of likely productivity. The next section of the paper presents evidence that the groups targeted by these programs are indeed stigmatized and that stigma generates systematic discrepancies between wage rates and realized productivity.

The Effects of Stigma on Government Training and Job Creation Programs.

The idea that being on welfare is stigmatizing has been around for a long time. Writing about direct relief recipients during the depression, W. E. Bakke wrote:

Now he has made a public declaration of his failure, and no rationalization can cover up the fact that a "reliever" is not among the roles his associates respect.¹²

Statistical proof that employer stigmatization of welfare recipients has major consequences has been harder to come by. In seeking such proof one needs to look for situations in which welfare recipient status is not confounded with some other variable (eg. completion of a training program) that is likely to affect either anticipated or real productivity. One also needs some way of quantifying employer beliefs regarding welfare recipients. Conceptually one wants to know the employer's willingness to pay to avoid hiring a welfare recipient. The establishment of a tax credit for hiring welfare recipients and other disadvantaged workers has created a number of opportunities for examining this issue.

The most compelling evidence on the stigma issue comes from two separate experiments conducted in 1980 in which welfare recipients were instructed to inform employers of their eligibility for tax credits and wage subsidies when they applied for a job. In both experiments the group that received these instructions had a lower placement rate than other eligible welfare recipients who had not. In the Dayton experiment, random assignment was used to select the groups and the reduction in the placement rate was statistically significant.¹³

The results of the Racine/Eau Claire, Wisconsin quasi-experiment are particularly interesting. The study compared WIN clients, who received

placement services prior to the initiation of the experiment, to clients served after the experiment began. Holding other characteristics constant, the WIN clients who were instructed to tell employers about their eligibility for a WIN tax credit were half as likely to obtain a job. This difference was statistically significant at the 10 percent level. A follow-up of a small subsample of experimentals and controls found that the WIN clients who followed instructions and brought up their WIN tax credit eligibility when contacting employers were one-third as likely to find a job. Of the thirty-two reporting that they used the WIN tax credit as a marketing tool, only two (6 percent) found jobs. Of the 26 reporting that they did not initiate discussion of the WIN tax credit eligibility, five (19 percent) found jobs.¹⁴ This difference is not statistically significant so it must be viewed as suggestive only.

The results of these experiments suggest that when welfare recipients announced their eligibility for a wage subsidy, prospective employers often gained information that stigmatized the applicant. Apparently for some employers, signaling one's welfare recipient status has such a powerful stigmatizing effect that the perceived reduction in the job applicant's expected productivity outweighs the prospect of a tax credit of up to \$3,000 in the first year and \$1,500 in the second year.

While welfare recipients are an important target group for JTPA services, most JTPA trainees are not on welfare. Are the JTPA clients who are not on welfare similarly stigmatized? Evidence that the disadvantaged in general - - not just welfare recipients -- are stigmatized is provided by a 1982 survey of 3000 employers regarding their use of and attitudes toward the Targeted Jobs Tax Credit (TJTC). During 1981-82 the TJTC offered employers a tax credit

equal to 50 percent of the first \$6000 of wages paid to newly hired eligible employees during their first year of employment and 25 percent of such wages paid during the second year of employment. The eligible target groups were handicapped individuals, welfare recipients and economically disadvantaged youth, Vietnam veterans, and ex-offenders. In that survey all employers who had heard of TJTC were asked if they thought "that tax-credit-eligible people usually make better or poorer new employees than people who are not tax-credit eligible." Even though the socially acceptable response is probably "don't know," "no difference," or "better," 28 percent of the respondents admitted to believing they were poorer than average. Only seven percent said they made better workers.

Further evidence of the important role of stigma comes from an examination of the effect of the TJTC on the hiring standards and wage offers of participating firms. Inducing employers to give hiring preference to TJTC eligibles was an important goal of the program. It was anticipated that many firms would be so attracted by the subsidy that they would be induced to lower their hiring standards in order to increase the number of TJTC hires. A simple way of tracking the effect of TJTC on a firm's hiring standards is to compare the wages and productivity of workers who were known to be TJTC eligible when hired to the wages and productivity of others hired for the same job. If TJTC causes firms to lower hiring standards, we would expect that for any given starting wage rate, new hires who were known to be eligible when hired would be less productive and require additional training.

If, however, being labeled as TJTC eligible is a bigger negative than the tax credit is a positive, then we would expect wage offers to be reduced or hiring standards (and therefore the productivity of the TJTC eligible new

hires) to be increased. The third possibility is that knowledge of who is TJTC eligible has no effect on hiring selections. This could occur if the stigmatizing effect of TJTC eligibility counterbalances the perceived benefits of the tax credit or if most firms are ignorant of the program or have no tax liability. Once they are created by the hiring process, discrepancies between productivity and wages can be expected to persist. As discussed previously, worker risk aversion, errors in measuring job performance, and the effective specificity of most productivity differentials between coworkers makes it optimal for employment contracts to allow only small adjustments of relative wage rates to differentials in perceived productivity.

A Model For Testing Hypotheses About Stigma

The profits exclusive of TJTC tax credit (Y_{1j}) generated by the "i"th new hire for the "j"th job are assumed to be equal to $(P_{1j} - T_{1j} - W_{1j})$, the productivity (P_{1j}) net of training (T_{1j}) and wage costs (W_{1j}) of that hire. Y_{1j} is a random variable which depends on the hiring standards and wage offers that were in effect when "i" was hired for job "j". Now let us compare the profitability of two recent hires, person 1 and person 2, by constructing a difference score, $(Y_{1j} - Y_{2j})$. Normally the expected value of the difference would be zero. If, however, the firm lowers hiring standards or increases wage offers in order to obtain workers who bring a subsidy to the firm, we would expect $(Y_{1j} - Y_{2j})$ to be positive when person 2, say, is TJTC eligible and when person 1 is not eligible. On the other hand, if knowledge that an applicant is TJTC eligible is so stigmatizing that it outweighs the attractions of the tax credit, we would expect $(Y_{1j} - Y_{2j})$ to be negative when person 2 is TJTC eligible and person 1 is not eligible. The difference between the

profitability of two new hires for the same job can, thus, be represented by the following equation:

$$Y_{1j} - Y_{2j} = B(S_{1j} - S_{2j}) + \Theta(H_{1j} - H_{2j}) + u_{1j} - u_{2j} \quad (1)$$

where S_{1j} = an indicator variable indicating whether person "i" was known to be eligible for TJTC when hired. (Thirty-one percent of the TJTC eligibles hired were not known to be eligible when hired and they are coded as zero on this variable.)

H_{1j} = a vector of control variables describing characteristics of the new hire known at the time of hiring (eg. work experience) and characteristics of the job such as hours worked and its temporary versus permanent character.

u_{1j} = a random error that is specific to individual i.

I generated the sample of jobs for which paired data are available in the following manner. A stratified random sample of employers was asked to provide information on "the last new employee your company hired prior to August 1981 regardless of whether that person is still employed by your company." They were also asked to provide data on a second new hire in the same or a very similar job but with contrasting amounts of vocational education. This resulted in data on 659 pairs of individuals who have the same or a very similar job at a particular establishment. (see Appendix A for a fuller description of the sample) Missing data on specific questions further reduces the sample. Seventy percent of the establishments had fewer than 50 employees and only 12 percent had more than 200.

The results of estimating equation 1 predicting the relative profitability of two new hires are presented in Table 1. Appendix A contains a complete

Table 1 about here

11

description of the construction of the dependent variable. Even though tax credit receipts are not part of the ex post profitability construct, new hires who were known to be eligible when hired turned out to be significantly more profitable for the firm both during the first quarter of employment and about a year later at the time of the interview (or separation if there is one).¹⁵ The magnitude of the gains enjoyed by firms appears to have been about 16 percent of the average productivity or wages of a worker with two years of tenure in the job. These results are consistent with the hypothesis of a pervasive stigmatization of TJTC eligibles. The employers who knew about TJTC eligibility prior to hiring (a) offered TJTC eligibles lower wage rates on average or (b) ended up getting employees of better than average productivity for that position. The stigma apparently caused employers to become extremely careful in their selection of TJTC eligibles. This care resulted in them hiring fewer eligibles but obtaining better employees than anticipated when they did hire an eligible. It would appear that some of the stigma of being TJTC eligible is not warranted.

The second line of the table presents an estimate of the effect of receiving a CETA on-the-job training subsidy on the profitability of a new hire. The coefficients on this variable are very close to zero implying that stigma effects roughly canceled out the attraction of the subsidy. Since firms must be familiar with the program to receive an on-the-job training contract and subsidies are paid in cash, ineligibility for the subsidy and ignorance of it cannot explain the lack of a negative coefficient. Stigma does not, however, seem to completely outweigh the subsidy's attractions as it does with TJTC.

Because training costs must be recovered more rapidly in temporary jobs, we would expect the measure of ex post profitability (which does not include amortization of training costs) to be larger for temporary jobs. This hypothesis is supported by the data. Hours worked per week appears to be positively related to profitability probably because hours can be adjusted upward when the new hire is discovered to be exceptionally competent. Profitability during the first quarter is positively related to relevant work experience and relevant vocational training and negatively related irrelevant work experience. By the time of the interview, relevant vocational training and irrelevant work experience no longer effects profitability and relevant work experience has developed a negative relationship. The issues raised by these results are left to another paper.

The Stigma Bias in Evaluations of Training Programs

It is very difficult to avoid stigmatizing the participants in training programs that are targeted on the disadvantaged. As the National Academy of Science's Committee on Youth Programs pointed out, "The problem is that this very targeting tends to create an image of the programs as designed only for 'failures;' both the programs themselves and their clientele become stigmatized in the process".¹⁶ Welfare recipients and low income youth respond to this fact by not mentioning their backgrounds when they apply for a job. Employers cannot ask job applicants whether they are on welfare or from a low income family because of legal proscriptions. As a result, most employed TJTC and JTPA eligibles are not known to be eligible by their employer. The Congressional Budget Office estimated that less than 10 percent of the employed disadvantaged youths eligible for TJTC had been claimed by their employers in 1981.¹⁷ When, however, disadvantaged youth complete a JTPA classroom

training program and seek employment, employers often learn of their disadvantaged status through their association with JTPA. The knowledge that the applicant is disadvantaged may hurt the individual's chances; the knowledge that they have been trained helps their chances. If classroom training is to benefit the trainee, the second effect must outweigh the first. If there are earnings gains for the trainees, they will probably be small.

Table 2 illustrates how stigma might bias evaluations of training programs serving the disadvantaged. The entries in this table are hypothetical employer perceptions of how productivity depends on training and on whether an individual is disadvantaged. Employers pay workers their expected productivity so the table also describes the wages offered new hires who signal these characteristics. The bottom row of the table is the average perceived productivity of all workers in the occupation (20 percent of whom are disadvantaged) and also the employer's wage offer if she does not know whether the job applicant is disadvantaged. This is what the disadvantaged individuals in the control group get paid. In this example, training raises productivity by \$2000, but when disadvantaged individuals are trained their wages rise only \$560 (\$8000-\$7440) because they become identified as disadvantaged by virtue of their training.

What happens to the remaining \$1440 of the productivity benefit of training? That depends on whether the perceptions that disadvantaged workers are less productive are correct. If these perceptions are wrong, then the employer who hires the trainee receives the \$1440 benefit because the firm gets a better employee than expected. This is apparently what occurred with TJTC. If the perceptions are correct on average (the disadvantaged individuals are indeed less productive), the employer who would have hired the

Table
2
about
here.

disadvantaged individual if they had not been trained is the beneficiary. That firm benefits because it avoids hiring an individual who is considerably less productive than expected and who would have received more in compensation than they contributed to the firm.

The conclusion is that training programs that have significant effects on a worker's literacy, skills and productivity may, because of this labeling effect, nevertheless, have very small effects on earnings. This may be the reason why the Job Corps which has produced reading gains as great as 2.5 months (in grade level terms) for every month of instruction, nevertheless, has very modest effects on employment and wages.¹⁸

Stigma may also help explain why comparison group studies have found that the most intensive form of CETA training, classroom training, had only modest effects on post-program earnings and why less intensive interventions such as public service employment and on-the-job training contracts had much more positive effects.¹⁹ It also helps explain why these studies find that work experience programs appear to have had the least favorable effect on their participants.²⁰ Community based organizations ran most work experience programs and a major share of classroom training projects. When program graduates sought unsubsidized jobs, prospective employers were able to tell from the name of the previous employer or training provider that the individual was disadvantaged. Being labeled in this way probably diminished the benefits of the training and work experience. Not all CETA trainees were publicly labeled, however. In the public service employment program and the on-the-job training contract program, sponsoring employers knew that the trainee was a CETA client but the later employers of these trainees generally did not. When these trainees completed their period of subsidized employment

and sought unsubsidized employment, their job experience at XYZ corporation or ABC agency served as a positive signal that gave no hint of their association with CETA.

The tendency of CETA participants in public service employment and on-the-job training to do better than work experience and classroom training participants is also a consequence of the mechanisms by which clients are recruited into these activities. The employers who participated in the public service and on-the-job training programs selected from among CETA eligibles. When private employers did not perceive the CETA referrals to be qualified, they hired an unsubsidized employee instead.²¹ Many of the CETA clients who were sent out to interview at potential contract employers were never offered a job and either drifted away from CETA altogether or transferred to other activities such as work experience and classroom training. (Currently in the agencies participating in the JTPA experiment, only about one-third of those initially sent in search of on-the-job training end up being formally enrolled as an on-the-job training client.) The providers of work experience and classroom training, on the other hand, were not generally able to pick and choose among CETA eligibles. Clients were formally enrolled in CETA only after they found a public service or work experience job or after beginning a classroom training program. Clearly this mode of recruiting eligible individuals into program activities causes selection bias in any evaluation based on "treatment" comparisons. Selection effects probably cause estimates of the effects of on-the-job training and public service treatments to be positively biased and estimates of the effects of work experience and classroom training to be negatively biased.

Stigma may also be the reason why, despite its very superficial nature, job search training has been found in experimental studies to have such positive effects on employment and earnings. Job clubs do not stigmatize participants because potential employers are approached directly by the job seeker, not by a job developer associated with a stigmatizing program. The programs train job seekers in interview techniques (eg. dressing appropriately, looking the employer in the eye when answering a question) which are designed to create an impression of high productivity. They do not teach skills that might actually increase productivity. If the productivity effects of job search training were known, they would probably turn out to be smaller than their earnings effects.

Local JTPA staff are quite aware of the labeling effect. They try to overcome it by promising local employers to screen their clients carefully and sending for interviews only those trainees who are able to handle the job. Some programs have succeeded in building a positive reputation and others have not.

Another way of forestalling stigma, and one that is recommended by the Committee on Youth Programs, is to "mainstream" the disadvantaged into institutions and activities that serve a cross section of the population.²² Financial aid programs such as Basic Educational Opportunity Grants and Work Study are successful in part because they avoid stigmatizing the students they aid. A mainstream institution, however, risks its reputation when it markets large numbers of disadvantaged workers to employers. A logical response to this problem is for the agency to disperse disadvantaged trainees across many institutions and fund them through performance based contracts. By making the reward both generous and contingent on placement, JTPA can induce

mainstream institutions to make the accommodations necessary for disadvantaged trainees to succeed in training and obtain jobs.

How to Avoid Stigma Bias in Future Evaluations

Clearly, if one wants to know the effect of training on productivity, one must ask those who supervise training program graduates (and control group members as well) about productivity. The ex post descriptions of the performance of specific named individuals are less likely to be distorted by prejudice (because they refer to specific individuals) and are better informed estimates of that individual's productivity than the starting wage offer. But are such reports truly comparable across jobs and across firms? Many dimensions of job performance (eg. absenteeism, tardiness, making suggestions for improvements in the business, willingness to work overtime, and responsiveness to training) are both concrete and comparable but global assessments of productivity are not. Supervisors are unable to compare their subordinates to similar employees in other firms. They are, however, able to make valid comparisons between subordinates who have the same job assignment, and they do this quite frequently.²³ What is needed is a measure of the discrepancy between an individual's productivity (net of required training costs) and wage rate which can be combined with the more conventional estimates of the earnings impacts of training.

The discrepancy between a worker's productivity and wage rate can be decomposed into 3 elements:

$$P_{1j} - W_{1j} = (P_{1j} - P_j) - (W_{1j} - W_j) + (P_j - W_j) \quad (2)$$

The first term is the worker's "relative productivity", the deviation of the "i"th worker's marginal revenue product net of required training costs (P_{1j}) from the marginal revenue product net of required training costs (P_j) of the

average incumbent in the job at the firm. The second term is the worker's "within-job relative wage", the deviation of an individual's wage from the mean for that job at the firm. The last term is the difference between the marginal revenue product net of required training costs of the average incumbent in the job (P_j) and the average wage for the job (W_j). Direct measurement of this last term is not feasible for it would require direct measures of the marginal revenue product of work groups that are comparable across jobs and across firms. Consequently, the effort to measure the discrepancy between MRP and wage rates should focus on the first two terms of equation 2.²⁴ There will no doubt be errors in measurement but the resulting estimate of the social benefits of training will be more valid than the estimates that result from assuming that $P_{1j}=W_{1j}$.

The best way to obtain assessments of $P_{1j}-P_j$ and $W_{1j}-W_j$ is to ask the supervisors of training program graduates to compare the trainee to two other employees with the same or a very similar job assignment. The first comparison would be against the most immediately senior employee (indicated by an s subscript). The second comparison would be against the most immediately junior employee (indicated by the r subscript). Data would be obtained on the tenure, wage, costs of required training and a variety of background characteristics of all three employees and on the productivity ratio for each of the two comparisons. In order to measure the effect of JTPA on relative productivity net of required training costs, similar data would be obtained for members of a randomly assigned control group. Productivity wage discrepancy ratios would be calculated from this data as follows:

$$DR_{1s_j} = A_j [P_{1j} - P_{sj} - (T_{1j} - T_{sj})] / P_j - (W_{1j} - W_{sj}) / W_j \quad (3)$$

$$DR_{1r_j} = A_j [P_{1j} - P_{rj} - (T_{1j} - T_{rj})] / P_j - (W_{1j} - W_{rj}) / W_j \quad (4)$$

where i indexes the trainee or control group member and P_j and W_j are the means for the three observations on workers in the job. A_j is a scaling factor equal to the ratio of value added to compensation in some industries and to one in others.²⁵ By estimating models predicting these discrepancy ratios as a function of tenure and trainee or control group status, we learn whether the trainee's productivity net of training cost has risen more or less than the wage differences between the two groups indicate.

If trainees' DR's are more positive than controls' DR's, the training is yielding social benefits -- higher productivity -- that do not accrue to the worker in higher wages, but instead go to employers, consumers, and other workers. If productivity net of required training costs is measured on a ratio scale and an assumption is made about A_j , the relationship between the average productivity and average wage of workers in the types of jobs occupied by trainees and their controls, an estimate of total social benefits of JTPA training may be obtained by combining the estimated mean DRs of the experimental group (DR_x) and the control group (DR_c) with the mean earnings of the two groups (Y_x , Y_c) in the following manner:

$$\text{Social Benefit} = (1 + DR_x)Y_x - (1 + DR_c)Y_c \quad (5)$$

Studies of productivity effects need not be expensive. Eight hundred phone interviews of supervisors split equally between treatment and control groups would yield an estimate of the mean difference between experimental and control group DR's with a standard error of .0242.²⁶ Such a study would have the power to detect very modest effects of training on relative productivity.

How important is it to know the social benefits of training programs like JTPA? It could be argued that the goal of JTPA is raising the earnings

of its disadvantaged clients and that benefits or losses incurred by other workers and employers should not be considered. This view is in error for three reasons. First, everyone's utility should enter the social welfare function. The fact that a particular program's primary objective is to aid the disadvantaged does not imply that the collateral benefits or costs experienced by employers, consumers, and other workers should be ignored. It only effects the weight to be attached to these effects.

Second, programs that are seen to benefit only a small and politically powerless minority tend to receive low levels of funding. If solid evidence were developed that employers, consumers, and other workers benefit from the training of the disadvantaged, political support for the programs would probably improve and program operators would find it easier to place trainees. Employer satisfaction with trainees is a central evaluation criterion for vocational education and this in part accounts for the political support that many employers give this institution.

Finally, the exclusive focus on earnings and placement rates in evaluations and in JTPA's performance standards tends to distort the management of the programs. It tends to induce creaming of the eligible population. It promotes the funding of short-term training in job search techniques that does not identify the client as being JTPA eligible at the expense of longer term training interventions that might increase a worker's productivity significantly. It tends to make public relations and placement quotas rather than high quality training the central focus of program management.

Summary and Caveats

I have challenged the widespread assumption that the wage effects of federal training initiatives are reliable and unbiased estimates of

productivity effects. This assumption has on occasion been justified by assertion that individuals are paid their individual marginal revenue product. This is clearly not true. Studies of the productivity of workers doing the same job and paid the same wage demonstrate that there is a great deal of variability in worker output.

Another way of justifying the equation of wage effects with productivity effects is to argue that wage offers reflect the reputation of individual training programs and that these reputations offer reliable and unbiased estimates of the value added of training. In order for this mechanism to work, employers must know that the individual was trained, and what the average effect of such training is on productivity at their work site. This assumption is not realistic. Most employers contacted by a JTPA graduate have never hired a graduate of that program before. Even where a few trainees have been hired in the past, the small numbers hired and the changing nature of the programs means that employers would not view this information as a reliable basis for predicting the success of future graduates. Consequently, employers incorporate reputational information gleaned from the mass media into their assessments. Because reputations only imperfectly correspond to reality, employer errors in assessing JTPA trainees are correlated. This implies that even though larger sample sizes can significantly improve the reliability of estimates of the wage effects of training, they may not substantially improve the reliability of estimates of the productivity effects of training.

Wage effects are not just unreliable estimates of productivity effects, they are biased as well. The people who participate in programs targeted on the disadvantaged are often stigmatized by this very participation. Welfare recipients and members of low income families who normally go incognito in

the labor market often become publicly identified by their participation. Because of this labeling effect, programs that substantially improve productivity may nevertheless have only modest effects on wage rates. Employers may end up benefiting as much from the training as the trainees. Analysis of the effect of TJTC on the hiring standards of participating firms supports the existence of a stigma bias.

Adding a few disclaimers to a study of the wage effects of training is not a sufficient response to the problems that have been identified. The wage numbers will get the attention; the disclaimers will be ignored. The proper response is to allocate resources to measuring the productivity effects of training when the analyst suspects the quality of training is not efficiently signaled or membership in a stigmatized group is revealed by training. Studies of productivity effects would supplement not displace studies of employment and wage effects.

I have argued that the standard practice of using the wage effects of training as estimators of productivity effects often results in a biased estimate of the social benefits of training. The paper has not shown that, for JTPA as a whole, the bias is large or that there are not other biases operating in the other direction. The magnitude of the bias is an empirical question that depends on the nature of the local labor market and the specifics of how the training program is administered and how trainees are marketed. It can be settled only by conducting studies in which the productivity effects are measured along with wage and employment effects.

ABOUT THE AUTHOR: John Hillman Bishop is an Associate Professor at the Center for Advanced Human Resource Studies, New York State School of Industrial and Labor Relations, Cornell University. This paper has been stimulated by my

membership on the Advisory Panel for the National JTPA Evaluation. The research reported here was supported by funds from Contract No. 99-4-576-77-091-01 of the Employment and Training Administration, U.S. Department of Labor. I would like to thank Kevin Hollenbeck, John Boudreau, Frank Schmidt, John Hunter, Robert Hutchens, John Gary and two unnamed referees for helpful comments on early versions of this paper. The opinions and conclusions expressed herein are solely those of the author and should not be construed as representing the opinions or policies of any agency of the United States Government.