Measuring Program Impacts on Earnings and Employment:

Do UI Wage Reports from Employers Agree with Surveys of Individuals?

Robert Kornfeld

Abt Associates Inc.

Howard S. Bloom

New York University

July, 1997

This paper is based on research supported under Contract Number 99-6-0803-77-068-01 with the U.S. Department of labor. It has benefited greatly from assistance provided by Larry L. Orr, Winston Lin, Stephen Bell, George Cave, Hans Bos, and an anonymous reviewer. All statements in the paper are those of the authors, however, and do not necessarily reflect the views of the U.S. Department of Labor or of persons from whom we received assistance.

Abstract

This paper attempts to determine whether wage records reported quarterly by employers to state Unemployment Insurance (UI) agencies provide a valid alternative to more costly retrospective sample surveys of individuals as the basis for measuring the labor market impacts of employment and training programs for low-income persons. This question is important because of the widespread and growing use of UI wage records for this purpose. We base our analysis on UI data and survey data for a sample of 12,318 low-income adults and out-of-school youths from 12 U.S. cities that were sites in the National Job-Training Partnership Act (JTPA) Study. The National JTPA Study was a Congressionally-mandated randomized experiment designed to measure the labor market impacts of JTPA programs for low-income persons. Our comparison of UI data and survey data indicates that for adult men, adult women, female youths, and all but a small subgroup of male youths, impact estimates based on UI data were comparable to those obtained from survey data, even though average earnings according to the survey data were somewhat higher than they were according to UI wage records. We therefore conclude that UI wage records can provide a valid alternative to surveys for measuring program impacts on many, but not all, groups of low-income persons.

Measuring Program Impacts on Earnings and Employment: Do UI Wage Reports from Employers Agree with Surveys of Individuals?

I. Introduction

Federal and state governments have funded many evaluations of programs to increase the earnings and employment of low-income persons. These studies of job-training programs, welfare reform initiatives, school-to-work transition activities, and other similar efforts, measure earnings and employment through one of two methods --- sample surveys of individuals or quarterly employer reports to state Unemployment Insurance (UI) agencies.

Due to the high cost of sample surveys and the increased availability and coverage of UI wage records, the latter approach has become increasingly popular. For example, virtually all evaluations of recent state-level welfare reform initiatives funded by the U.S. Department of Health and Human Services relied at least partly on UI wage records.[1] In addition, the U.S. Department of Labor, the U.S. Department of Agriculture, and other federal agencies have used UI wage records to evaluate a wide range of job-training programs and proposed changes to the Unemployment Insurance system.[2] UI data also have been used for administrative purposes, to monitor the performance of job-training programs, to identify earnings in order to determine child

[1] For example, the numerous studies reviewed by Gueron and Pauly (1991), plus later studies of Job Opportunities and Basic Skills (JOBS) programs in Ohio (Fein, Beecroft, and Blomquist, 1994), California (Riccio, Friedlander, and Friedman, 1994), Florida (Kemple and Haimson, 1994), Michigan (Werner and Kornfeld, 1996) used UI wage records. While most recent HHS-sponsored analyses of changes in the welfare system used UI wage records, only a few, such as the New York State Child Assistance Program (Hamilton, et. al., 1993) and an earlier study, the AFDC Homemaker-Health Aide Demonstrations (Bell and Orr, 1994) used survey data to measure earnings.

[2] UI wage records were used to study the effects of reemployment bonuses for UI claimants in New Jersey (Corson et. al., 1989) and Pennsylvania (Corson et. al., 1991). They were also used to study the effects of job-search assistance for UI claimants (e.g., see Benus, et. al., 1996 and Bloom, 1990). For a summary of these and other studies that used UI wage records to study the impacts of programs for UI claimants, see Meyer (1995).

support payments, and to verify the earnings of welfare recipients.[3]

To help assess the appropriateness of substituting UI wage records for surveys of individuals in future program evaluations, we compare impact estimates obtained from both data-sources for the same group of low-income persons. To the extent that the two data-sources provide similar findings, UI wage records might provide a viable alternative to surveys.[4]

Our analysis is based on earnings and employment data from the National JTPA Study, the Congressionally-mandated evaluation of Title II-A of the 1982 Job Training Partnership Act.[5] JTPA Title II-A is the nation's employment and training program for low income adults and out-of-school youths with significant barriers to employment. Funds for this program are distributed to over 600 local service delivery areas (SDA's), whose staff plan and implement programs with guidance from a local Private Industry Council (PIC).[6] Such programs usually include a mix of classroom training in occupational skills, on-the-job-training, job search assistance, basic education, work experience, and other services.[7]

Members of the 21,000-person National JTPA Study sample applied to the 16 local SDAs who took part in the study between November, 1987 and September, 1989. SDA staff

---

[3] For a discussion of these applications, see OTA (1994).

[4] If findings from the two data-sources differ, one would need a benchmark measure of "true" earnings in order to choose between them. Such a benchmark was not available for the present analysis, although we used a third data-source, earnings reported by employers to the U.S. Internal Revenue Service, as a basis for some comparisons.

[5] See Bloom *et. al.* (1997) and Orr *et. al*. (1996) for a detailed description of the study.

[6] PICs are comprised of representatives from businesses, unions, social service groups, employment and training agencies, and other local organizations.

[7] JTPA services can be provided by a wide range of institutions, including high schools, community colleges, vocational schools, and community-based organizations.

then recommended eligible applicants for one of three broad service strategies: (1) classroom training, (2) on-the-job-training/job-search assistance, and (3) miscellaneous services (which focused mainly on immediate employment for adults and basic education for youths). Applicants within each of these "service strategy subgroups" were then randomly assigned to a program group or a control group, with two program group members for each control group member.

Program group members could receive JTPA services.[8] Control group members were not able to receive JTPA services, but could receive other local services for which they were eligible.[9] Estimates of program impacts then were obtained by comparing the earnings and employment of program group members with those of control group members. Impact findings were reported for five "target groups" -- adult men, adult women, female youths, male youths with no prior arrests, and male youths with prior arrests.[10]

The present paper compares estimates of program impacts on employment and earnings for the **12,318 sample members from the 12 study sites for which data from both sources were available**. The two data-sources produced roughly the same impact findings for four of the five study target groups. Hence, for these groups, UI wage records provide a valid alternative to sample surveys. However, for the small subgroup of male youths with past arrests, survey-based impact estimates differed markedly from those based on UI wage records. Because it was not possible to fully identify the sources of this discrepancy, we could reach no firm conclusion

[8] Although program group members could opt for services outside their recommended service strategy, these strategies were strongly correlated with services actually received.

[9] Control group members received much less service than program group members.

[10] Orr, *et. al.* (1996) explain why male youths were separated into two subgroups.

about which data-source was most accurate for the subgroup.

The remainder of this paper: (1) describes how we constructed the survey and UI measures of employment and earnings; (2) compares estimates of program impacts and measures of employment and earnings obtained from the two data-sources for the same sample of individuals; (3) reports the results of tests designed to identify specific problems in each data-source; and (4) summarizes our findings and discusses their implications.

II. The Data

The data for this study come mainly from two consecutive sample surveys which requested self-reported retrospective data on earnings, and from UI wage records which provide earnings data for an overlapping sample of persons and calendar quarters. For comparative purposes, quarterly earnings are analyzed because the UI wage records provide data in this form only.

Self-reported earnings data from two sample surveys of individuals. The period of coverage of these surveys varied across sample members. The First Follow-up Survey, which was conducted between October, 1989, and January, 1991, interviewed 17,217 JTPA Study sample members. The survey covered the time between the dates of random assignment and the interview, a period ranging from 12 to 37 months, with an average period of 21 months. The Second Follow-up Survey, which was conducted between July and December of 1991, interviewed a random subset of 5,468 members of the initial JTPA study sample. For most respondents, the Second Follow-up Survey focused on events between the interview and the time

of the First Follow-up Survey.[11] This second interview was conducted 20 to 48 months after random assignment (35 months, on average.)[12]

The surveys asked respondents to report the start and end dates of all job spells which fell within the period covered, and then to provide earnings data for each spell. Respondents paid on an hourly basis reported their non-overtime hourly wage at the start and end of the job spell. Those paid daily, weekly, every two weeks, twice a month, or monthly reported the amount they usually earned per period (excluding overtime), and whether the reported amount was gross or net of taxes. Those paid for different periods reported the number of times they were paid over the job spell and the usual amount paid per period. These data (and, when needed, data on hours worked) were used to calculate quarterly regular earnings.

For each spell, respondents also reported earnings from overtime, tips, bonuses, and commissions. Respondents who said they received paid overtime were asked "How much overtime pay did you usually earn in an average week, before taxes and other deductions?" When answering this question, respondents probably reported their average overtime pay for weeks in which they worked overtime, instead of the overall average for all weeks in the job.[13] Thus, measured overtime earnings probably overstate actual overtime earnings. To deal with this issue,

---

[11] For the 472 respondents who had not answered the First Follow-up Survey, the Second Follow-up Survey focused on events between random assignment and the Second Follow-up interview.

[12] The First Follow-up survey achieved a response rate of 83 percent. The Second Follow-up survey achieved a response rate of 78 percent.

[13] It would have been nearly impossible for respondents to report average overtime earnings for all weeks worked during a job spell, because to do so would have required them to mentally: (1) compute their number of weeks with and without overtime and (2) take a weighted mean of their average overtime pay for weeks when it was earned and zero overtime pay for weeks when it was not earned. Because respondents were not asked how many weeks they had worked overtime, it was not possible to make this calculation for them after the fact.

survey earnings measures used in the present paper include only half of the overtime earnings reported by survey respondents. This adjustment was based on a detailed analysis of survey/UI earnings ratios for person-quarters with and without overtime earnings by Orr et. al. (1994).[14] The survey also provided data on earnings from tips, bonuses, and commissions. For each spell identified, respondents who said they received these payments provided the amount earned in a typical week.

This study analyzes total quarterly earnings, defined as the sum of regular earnings, and earnings from overtime and from tips, bonuses, and commissions. For dual and multiple job holders who report each of these jobs, our total earnings measure sums earnings from all jobs within each quarter. Respondents were also asked to report total earnings from "odd jobs" not mentioned as spells. To calculate total quarterly earnings, the earnings from these odd jobs were divided evenly over all quarters in the covered period. The analysis omitted quarterly observations with any missing or imputed data (including outliers)[15], and observations of quarters the survey covered only partially because the interview date occurred midway through a quarter.

The two surveys provided other potentially useful information. Respondents reported whether there were any weeks during each spell in which they were not paid, although this information was not used to estimate earnings. For the most recent reported spell, respondents provided information on the job characteristics -- the type of employer (government, private,

[14] We included half of overtime earnings for all target groups, which is a simplification of the approach used by Orr et. al. (1994), pp. 202-203, who included: 35 percent of overtime earnings for adult women, 70 percent for adult men, 42 percent for female youths, 50 percent for male youths with no prior arrest and 76 percent for male youths with prior arrests.

[15] About 8 percent of person-months of data were missing or imputed. See Orr *et al* (1996).

military, or self-employed), industry and occupation, and a verbal job description.

Earnings data from State UI Wage Records UI wage records consist of total quarterly earnings reported by employers to state UI agencies for each employee. By law, any employer paying $1,500 in wages during a calendar quarter to one or more employees is subject to a state UI tax, and hence, must report quarterly what is paid to each employee, including regular earnings, overtime, and tips and bonuses.[16] However, there are some important exemptions. UI wage records do not cover self-employed persons, military personnel, federal government workers, railroad employees, some part-time employees of non-profit institutions, employees of religious orders, some students employed by their schools, and most independent contractors. Any state's wage records will miss earnings from out-of-state jobs, from persons with incorrectly recorded social security numbers, or from employers who fail to report covered earnings.

State UI administrators in 12 states agreed to provide all wage records for the period of the study for all sample members in the state, using social security numbers to identify these records. Persons with multiple employers have separate wage records from each employer, so earnings from multiple jobs may be summed to compute total earnings per quarter. Persons for whom no wage records exist in a quarter are assumed to have zero earnings (are not employed) during the quarter. This study omitted observations which were classed as outliers or which came from incomplete quarters.[17]

[16] Agricultural employers must report earnings if they have either a quarterly payroll of at least $20,000 or have hired 10 or more employees in each of 20 or more weeks during the preceding calendar year. Employers of paid household help must report wages if they pay at least $1,000 in cash wages during any quarter.

[17] Because employers may need a few months to report these earnings, some wage records provided very soon after a calendar quarter were thought to be incomplete, and excluded from this study.

Samples of persons and quarters used in this study. This study compares total quarterly earnings reported by these two data sources, for the same calendar quarters and the same persons. Observations of "person-quarters" are included in our sample only if they have non-missing (non-imputed, complete) quarterly earnings from both sources. Person-quarters in our sample may have zero earnings reported by one or both sources.

Our analysis is therefore based on the post-random assignment experiences of 12,318 sample members from 12 study sites, and includes 70,572 person-quarters. Because sample members could respond to one or two surveys, and because of variation in survey coverage dates, instances of missing data, and numbers of quarters of available UI wage records, the number of usable quarters of data per person varies between 4 and 9 quarters, with an average of 7 quarters per person. Because the sample includes only person-quarters with non-missing data from both sources, the problem of survey non-reponse is not considered here.

III. Impact Estimates from the two Data-Sources

This section compares estimates from survey data and UI data of JTPA impacts on: (1) mean quarterly earnings of the five JTPA Study target groups, (2) their quarterly employment rates, and (3) mean quarterly earnings for sample members who were AFDC recipients or UI claimants when they applied to JTPA.

Impacts on Mean Quarterly Earnings by Target Group

Exhibit 1 summarizes our main impact findings for earnings. Its first column presents results from the follow-up surveys and its second column presents corresponding results from UI wage records. For example, survey responses indicate that adult women in the program group earned $1,294 per follow-up quarter, on average, whereas adult women in the control group earned $1,141 per quarter. The statistically significant difference of $153 per quarter, provides a survey-based estimate of the JTPA impact.[18] When expressed as a percentage of control group earnings, this difference implies a 13.4 percent program-induced earnings gain.

Corresponding findings from UI wage records are $1,048 for the program group, $922 for the control group, and thus, a statistically significant $126, or 13.7 percent program-induced earnings gain. Hence, even though survey earnings were somewhat higher than UI earnings, on average, their impact estimates were roughly comparable in dollars, and virtually identical in percentage terms. The two impact estimates are similar because the ratio of mean survey earnings to mean UI earnings was almost the same for the program and control groups (1.23 and 1.24, respectively; see column three). This ratio was computed as follows:

$$\frac{\sum_i \sum_t S_{it} / n}{\sum_i \sum_t U_{it} / n} = \frac{\overline{S}}{\overline{U}}$$

[18] The statistical significance of this difference represents the statistical significance of the impact estimate.

S_{it} = *Survey-reported earnings for person i during quarter t*
U_{it} = *UI-reported earnings for person i during quarter t*
n = *Total Number of person-quarters.*

If the survey/UI earnings ratio is the same for the program and control groups, it is as if the two data-sources reported earnings in different scales that were proportional to each other. Consequently, their impact estimates will be "scaled" in the same proportion. For example, if survey earnings were exactly 25 percent higher than UI earnings for the program group and the control group, their difference in survey earnings will be exactly 25 percent higher than their difference in UI earnings. Hence, the dollar impact estimate obtained from survey data will be 25 percent higher than its counterpart from UI data. However, when this impact is expressed as a percentage of average control group earnings, the findings from both data-sources will be identical.[19] This could occur if the mechanisms that produced the differences between the two data-sources were the same for program and control group members, which might be the case if the program did not markedly affect the types of jobs taken by sample members.

---

[19] The argument is as follows.

If: $Y_p = "Y_p^*$
and $Y_c = "Y_c^*$
where:
Y_p and Y_c = mean survey earnings for the program group and control group, respectively
Y_p^* and Y_c^* = mean UI earnings for the program group and control group, respectively
then:
$Y_p - Y_c = "(Y_p^* - Y_c^*)$
and
$(Y_p - Y_c)/Y_c = " (Y_p^* - Y_c^*)/"Y_c^*$

$= (Y_p^* - Y_c^*)/Y_c^*$

As for adult men, the survey/UI earnings ratio is also almost identical for the program group and the control group (1.32 versus 1.30, respectively). Consequently, their quarterly impact estimates in dollars are similar (a statistically insignificant $94 versus a statistically insignificant $59), and their impact estimates in percent are almost the same (5.2 percent versus 4.2 percent).[20] For adult men and women then, UI wage records and individual sample surveys produced JTPA impact estimates that "tell the same story". In fact, differences between their estimates were much too small to affect the benefit/cost conclusions of the National JTPA Study (see Orr *et. al.*, 1996).

Survey and UI data also provided consistent impact estimates for female youths and male youths without a prior arrest. Both data-sources indicated that JTPA had no impact for female youths ($2 versus $1 per quarter or 0.0 percent versus 0.0 percent); and a small, statistically insignificant negative impact for male youths without a prior arrest (-$99 versus -$88 per quarter or -6.0 percent versus -8.0 percent). The survey/UI earnings ratios were virtually identical for their program and control groups.

In contrast, the two data-sources produced markedly different impact estimates for the one-out-of-four male youths with a prior arrest. Survey data suggest that JTPA sharply reduced their future earnings (by a statistically significant - $248 per quarter or - 16.2 percent); whereas UI data suggest that it had no effect (- $1 per quarter or - 0.0 percent).[21] This discrepancy arises

[20] One can examine the benefit-cost implications of these differences through a sensitivity analysis which asks whether the two data-sources produce "bottom-lines" which differ materially.

[21] The policy implications of these two different findings are the same, however --- neither suggests that JTPA was successful for male youths with a prior criminal record.

from the fact that the survey/UI earnings ratio for program group members was much lower than that for control group members (1.69 versus 2.01). We return to this issue below.

Impacts on Quarterly Employment Rates by Target Group

Exhibit 2 compares estimates of JTPA impacts on quarterly employment rates obtained from survey and UI data. Quarterly employment rates were defined as the percentage of person-quarters for which non-zero earnings are reported. Hence, they only indicate whether any employment occurred during a quarter. They do not indicate how much employment occurred (how many weeks, days or hours were worked). Employers do not report this more precise information to state UI agencies.

As can be seen, there was little difference between quarterly employment rates computed from the survey and UI data; the survey/UI ratio was very close to one for all target groups. Consequently, there was little difference between the employment impact estimates obtained from the two data sources. Because employment rates are an important outcome measure for job-training programs, it is encouraging to note that relatively inexpensive UI data may provide an acceptable substitute for survey-based estimates of impacts on this outcome.

Exhibit 2 also shows that survey/UI differences in quarterly employment rates are much smaller than survey/IU differences in average quarterly earnings. Hence, the observed differences in reported earnings must reflect either differences in the amount of work recorded for each quarter employed, differences in the rate of pay recorded for this work, or both.

Impacts on Quarterly Earnings for AFDC Recipients and UI Claimants

UI wage records have been used to estimate the impacts of major welfare reform initiatives (for example, Werner and Kornfeld, 1997; Fein, Beecroft and Blomquist, 1994; Riccio, Friedlander and Friedman, 1994; Kemple and Haimson, 1994) and large-scale UI demonstration programs (for example, Bloom, 1990, Corson *et. al.* , 1989, Corson, *et. al.*, 1991, Spiegelman *et. al.* 1992). We therefore compared program impact estimates from UI data and survey data for AFDC recipients and UI claimants in the National JTPA Study sample. Exhibit 3 presents our findings for adult women and female youths who were receiving AFDC benefits when they entered JTPA, plus adult men and adult women who were receiving UI benefits when they entered JTPA.[22]

For adult women receiving AFDC (the largest subgroup by far in the exhibit) survey and UI data produced very similar percentage impact estimates (26 percent and 24 percent, respectively) because the survey/UI earnings ratio was almost the same for their program and control groups (1.27 and 1.24, respectively). The corresponding dollar amounts of these impact estimates were thus, "scaled" accordingly ($208 from the surveys versus $153 from the UI data). Regardless of the data-source used, however, the impact story for women on AFDC was the same -- JTPA markedly increased their earnings.

Survey and UI data also produced consistent stories for female youths receiving AFDC

[22]We do not report findings for AFDC recipients or UI claimants in the other JTPA Study target groups because these subgroups were too small to do so (they each comprised less than 200 persons).

and adults receiving UI; there was no statistically significant impact estimate from either data-source for any of these subgroups. Point estimates from the two data-sources differed somewhat, however, because the survey/UI earnings ratio for the program and control groups differed. Nevertheless, the differences in impact estimates from the two data-sources were not statistically significant, probably because of the small samples involved and the large amount of random error in measures of individual earnings.

IV. Comparing Earnings Measures from the Two Data-Sources

This section begins our search for explanations of the observed survey/UI difference in earnings. It focuses on the extent to which the overall survey/UI difference reflects large differences for a few sites, large differences for a few persons, and differences in the average earnings of jobs in quarters with earnings reported by only one data-source. It also focuses on the extent to which random versus systematic error produced the survey/UI earnings differences for individual sample members.

Survey/UI Earnings Comparisons Across Sites

Exhibit 4 indicates that survey earnings are higher than UI earnings for every target group from every study site. The exhibit also indicates that the overall survey/UI discrepancy cannot be explained by a few sites with unusually large discrepancies. Furthermore, tests conducted, but not reported here, indicate that estimates of JTPA impacts on earnings from the two data-sources are not highly sensitive to the exclusion of any particular site. Hence, the mechanisms which

produced the survey/UI discrepancy exist in all of the JTPA Study sites.

Survey/UI Earnings Comparisons Across Individuals

Exhibit 5 presents the distribution of survey/UI earnings differences**,** in dollars, for individual sample members.[23] As can be seen, these individual-level differences frequently were quite large (many were above $1,000, for example). Furthermore, their magnitude and direction varied widely across sample members. Because survey/UI earnings differences reflect substantial random measurement error in one or both data-sources (discussed below), there appears to be a considerable amount of such error at the individual level which is not visible when mean earnings are compared.

The bottom panel in the exhibit compares the survey/UI earnings ratio for the full sample for each target group with its counterpart excluding the largest 2.5 percent of the positive survey/UI earnings differences and the largest 2.5 percent of the negative differences. By "trimming" the distribution in this way it is possible to examine the effect of unusual individuals on the overall findings. As can be seen, there was no such effect.

Random and Systematic Error in the Earnings Measures

Exhibit 6 separates individual-level survey/UI earnings differences into their random and systematic components in order to determine the mix of random and systematic measurement

[23] For the 3.8 percent to 14.3 percent of sample members with no survey/UI difference in mean quarterly earnings, both data-sources reported zero earnings.

error that produced these differences. The first row in the exhibit presents the root mean squared individual difference --- the average size of this difference, regardless of its direction. The second row presents mean UI earnings. As can be seen, individual errors are quite large relative to average earnings.[24]

The bottom of the exhibit indicates that 76.2 percent to 93.2 percent of individual survey/UI earnings differences are due to random measurement error. This random error reduces the reliability of program impact estimates but does not impart a bias to them.[25] The remaining and much smaller systematic error component will not affect the reliability of program impact estimates, but it can impart a bias to them. However, if systematic error affects program and control group members similarly, it will not affect percentage impact estimates.

The Relationship Between Reported Earnings and Reported Employment

To examine further the differences between survey and UI earnings measures, Exhibit 7 compares them under three different conditions: (1) for person-quarters in which only the surveys reported employment (column one); (2) for person-quarters in which only UI data

[24] It does not matter whether we use mean quarterly UI earnings or mean quarterly survey earnings for this purpose, since both are roughly comparable.

[25] The decomposition of RMSD was based on the following relationship:

$$MSD = VAR(D) + MD^2$$

where:

- MSD = $RMSD^2$ (a summary measure of total individual differences),
- VAR(D) = the variance of individual differences (a summary measure of the random component of individual differences), and
- MD^2 = the mean individual difference squared (a summary measure of the systematic component of individual differences).

VAR(D) was 96 percent of RMSD for adult women, 94 percent for adult men, and so on.

reported employment (column two); and (3) for person-quarters in which both data-sources reported employment (columns four and five).[26]

When only the surveys reported employment (and UI data presumably missed it), mean earnings were more than twice what they were when only UI data reported employment (and the surveys presumably missed it).[27] This suggests that surveys are more likely to miss "low-earnings" quarters, perhaps because respondents forget about minor, or short-term jobs. In contrast, UI data appear more likely to miss "average earnings" quarters --- where mean earnings are similar to that when both data-sources report employment. This might be due to random errors in matching UI wage records, out-of-state jobs, jobs that are not covered by UI, and/or earnings that are "off the books".

Even when both data-sources reported employment, however, survey earnings were higher than UI earnings. This suggests that the two data-sources might occasionally report earnings for different jobs, and/or different earnings for the same jobs.

V. Empirical Tests of Each Data-Source

To explore some of the hypotheses generated by the preceding findings, we next report the results of tests for specific problems in each data-source. We consider UI wage records first

[26] Note that in Exhibit 7, person-quarters for program group members and control group members were pooled for each target group.

[27] Without knowledge of "true" earnings, there is no way to validate the presumption that when only one data-source reported employment it was missed by the other. Nevertheless, we believe that this is a plausible assumption to make and that the implications which flow from it are plausible.

and follow-up surveys next.

Tests of UI Earnings Measures

As suggested above, some types of jobs are not covered by Unemployment Insurance and thus, are not reported to state UI agencies (well over 90 percent of all jobs are covered in most states, however); and some jobs are located out-of-state and thus are not reported to the UI agency in a sample member's state of residence. In addition, some earnings for in-state jobs that should be covered by UI are not reported by employers. For example, they might accidently fail to report the earnings of some employees, or might under-report earnings, especially for persons in short-term, low-wage jobs. Furthermore, employers might deliberately fail to report some wages in order to avoid paying Unemployment Insurance taxes, to avoid a later responsibility for Unemployment Insurance benefits, or to cooperate with employees who wish to conceal their earnings. In addition, UI data might miss wages that are reported with incorrect Social Security numbers.

Measuring the *Combined Effect* of Uncovered Jobs, Out-of-State Jobs and Unreported Earnings by Comparing UI Earnings with IRS Earnings: Unlike UI wage records, earnings reported by employers to the Internal Revenue Service (IRS) should include jobs that are not covered by UI and jobs that are located out-of-state. In addition, there are financial incentives for employers to *under-report* earnings to state UI programs and *fully-report* earnings to the IRS.[28]

Earnings reported by employers to state UI agencies provide the basis for assessing a payroll tax that finances UI benefit payments. In contrast, earnings reported by employers to the

[28] Our thanks to Larry Orr for bringing this point to our attention.

IRS are a business expense that can be deducted from a firm's earnings and thereby will lower its income tax payments. One thus, can measure the combined effect on UI earnings of uncovered jobs, out-of-state jobs, and unreported earnings by comparing mean earnings from UI wage records to mean earnings from IRS records (employer reports).

IRS-earnings are reported by calendar year and were made available in the form of means for groups of 10 to 19 persons to protect individual confidentiality. We were able to align some of our quarterly UI-earnings records with some of the annual grouped IRS- earnings records to compare mean earnings for each group from each data-source.[29] Exhibit 8 presents these findings. As can be seen, the ratio of mean IRS-earnings to mean UI-earnings ranged from 1.14 for adult women to 1.25 for male youths.[30] In other words, IRS earnings were 14 percent to 25 percent higher than UI earnings, suggesting that UI wage records are missing earnings from some jobs.

A comparison of the IRS/UI earnings ratios in Exhibit 8 with corresponding survey/UI earnings ratios in Exhibit 1, suggests that **about half of the survey/UI earnings difference reflects earnings that are missing from UI wage records**. For example, the survey/UI earnings ratio for adult women was about 1.23 (Exhibit 1), whereas the corresponding IRS/UI ratio was 1.14 (Exhibit 8). This suggests that 14 points of the 23 percentage-point gap between

[29] Because the JTPA follow-up surveys were administered at different times of the year, much of the survey data covered portions of years which could not be matched to a full calendar year of IRS earnings data. Furthermore, the second follow-up survey was administered to only a fraction of the study sample. Hence, there were not many IRS groups in which all 10 to 19 sample members had a full calendar year of survey earnings data to match with corresponding IRS earnings data. Consequently, we could not align the IRS data with the survey data for a sample that was large enough to support a reliable comparison of these two data sources.

[30] Because the IRS groups for which data were obtained did not separate male youths with prior arrests from male youths without prior arrests, it was not possible to separate these two subgroups for this part of the analysis.

survey-earnings and UI-earnings are due to earnings that were not reported to UI.[31] Similar findings were obtained for the other JTPA Study target groups.

Measuring the *Separate* Effect of Out-of-State Jobs by Comparing the Survey/UI Earnings Ratio for Sites that Were Located Near a State Border with The Ratio for Other Sites. To the extent that out-of-state jobs were an important source of earnings for some of our sample members, the survey/UI earnings ratio should be higher for the seven study sites that were located near a state border than for the five sites that were not.[32] However, there was virtually no difference between the survey/UI earnings ratios for these two groups of sites. Thus, out-of-state jobs do not explain why UI wage records reported lower earnings than sample surveys.

Measuring the *Separate* Effect of Uncovered Jobs by Using the Survey to Identify These Jobs. Follow-up surveys for the National JTPA Study collected detailed information about the nature of each respondent's most recent job. Because roughly two-thirds of the person-quarters in our sample involved a most recent job, it was possible to determine whether or not these jobs were in categories that typically are not covered by Unemployment Insurance (for example, self-employment, a job with one's family, a military job, or a job with the federal government).

For person-quarters when earnings are reported by the survey and by UI wage records, these types of jobs are not be likely to be present. For person-quarters when earnings are reported only by the survey (and thus, are potentially missed by UI records), these types of jobs

[31] Because it was necessary to use different samples for each analysis in this section, their findings are not entirely comparable to each other or to the survey/UI ratios in Exhibit 8. Nevertheless, the stability of the basic findings across sites and the clear pattern of findings presented in this section provide support for the interpretation that we give them.

[32] A site was classified as being near a state border if it included any counties that were adjacent to another state.

are more likely to be present.

Consistent with this expectation, only *2.6 percent* of the person-quarters with non-zero survey- and UI-earnings involved what appear to be uncovered jobs. In contrast, *15.0 percent* of the person-quarters for which only the survey reported non-zero earnings involved these types of jobs. Therefore, uncovered jobs probably account for part of the difference between survey-earnings and UI-earnings. However, if only 15.0 percent of the person-quarters that appear to be missing UI-earnings involve an uncovered job, this phenomenon can only account for a small fraction of the survey/UI earnings difference.

Implications of the Preceding Findings. It appears that *some combination* of uncovered jobs, out-of-state jobs, and unreported earnings accounts for about half of the survey/UI earnings difference. It also appears that out-of-state jobs did not affect UI wage records and that uncovered jobs had only a small effect.[33] By process of elimination then, it would seem that **under-reporting of earnings or non-reporting of certain jobs explain almost half of the observed survey/UI difference**.[34]

If JTPA has only a modest impact on the labor market behavior of its participants -- which appears to be the case -- then there is little margin for it to influence the types of jobs they hold. This, in turn, suggests that under-reporting or non-reporting of earnings by employers

[33] Another potential source of problems with UI wage records is that some employers might have submitted earnings records that were too late for inclusion in our analysis. This, in turn, would have caused UI wage records to further understate earnings. If our data were subject to such a reporting lag, the survey/UI earnings ratios for our most recent person-quarters would be higher than corresponding ratios for less recent person-quarters. This was not the case, however, because states continually updated their UI wage records to include previously late employer submissions.

[34] This conclusion must be qualified, however, because the three analyses upon which it is based reflect different samples.

to state UI agencies (which probably depends on the type of job one has) would be similar for program group members and control group members.

Tests of the Survey Earnings Measures

A number of problems can arise when measuring earnings from a survey. For example, respondents might fail to recall certain jobs, especially brief, informal ones. In addition, they might deliberately conceal some earnings to avoid losing welfare or other transfer payments, to escape child support payments, to escape detection as illegal immigrants, or to conceal illegal income. Conversely, survey respondents might falsely report earnings from nonexistent jobs or overstate earnings from actual jobs to exaggerate their success. Errors in the reported timing of jobs (their start dates and end dates) also could produce errors in reported earnings.[35]

Testing for Survey Recall Bias by Comparing Earnings Measures as a Function of the Recall Period. The accuracy of an earnings measure obtained from a survey depends on the ability of survey respondents to recall past events. Random survey recall error will reduce the reliability of a survey earnings measure but it will not bias the measure. Systematic recall error (consistent over- or under-reporting) will bias the measure and thus could affect the survey/UI earnings ratio.

Recall error probably most often involves jobs that are forgotten and hence, earnings that are understated. The potential for such recall error probably increases as the length of the recall

[35] Survey *nonresponse* can also bias program impact estimates, but we do not address this issue. We only compare survey and UI earnings and employment for individuals with data from both sources.

period increases. To the extent that recall error exists in our survey data, we thus should expect the survey/UI earnings ratio to decrease as the recall period increases. We found, however, that the survey/UI earnings ratio was stable over widely varying recall periods. Survey reports of earnings during the past several months (which should have little potential for recall bias) were about as accurate, relative to UI earnings, as were survey reports for 19 or more months in the past. Hence, there was no temporal pattern suggesting that recall bias affected our survey-based earnings measures.[36]

Testing for Exaggerated Self-Reported Earnings by Examining the Extent to Which Survey Responses Indicated Unusually Good Jobs. Some survey respondents might have overstated their earnings in order to impress survey interviewers. To test for this possibility, we examined detailed job descriptions that were obtained from the survey for the most recent job held by each respondent plus data on employer names from UI wage records obtained from several states in our sample. Two types of analyses were conducted.

One analysis focused on person-quarters for which detailed job information was available from both the survey and UI wage records. For this analysis, we checked to see if the types of jobs reported by the two data-sources were systematically different. Although there were numerous apparently random differences, there was no evidence that the surveys reported jobs that were noticeably "better" than those reported by UI wage records.

A second analysis focused on person-quarters where the discrepancy between survey earnings and UI earnings was large, and detailed job information was available from the survey.

[36] This analysis cannot, however, identify survey recall bias that is independent of the recall period.

Inspection of the jobs reported by the survey suggested nothing unusual about them. There were few, if any, implausibly "good jobs". The jobs reported by survey respondents were mainly low-wage jobs in settings that are typical of such jobs, such as fast food restaurants, local retailers and small manufacturing firms.[37]

Testing for Other Features of the Survey Responses that Might Help Explain the Survey/UI Discrepancy. As discussed in Section II, survey-reported overtime earnings were reduced by 50 percent to adjust for their apparent exaggeration, but this cannot account for the full survey/UI earnings discrepancy. Even without overtime earnings, the survey/UI earnings ratios are 1.18 for adult women, 1.22 for adult men, 1.29 for female youths, 1.42 for male youths with no prior arrest, and 1.67 for male youths with a prior arrest.[38]

The survey/UI discrepancy also cannot be explained by unusual pay periods or by confusion over pre-tax versus after-tax earnings. Only 1-2 percent of the person-quarters with survey-earnings had earnings that were reported on a non-standard basis, or reported after-taxes.[39]

Further tests indicated that the survey/UI earnings discrepancy was not related to whether earnings were reported from tips, bonuses, commissions, or Aodd jobs@, or whether

[37] Our methodology could only detect "big lies"; it could not detect minor exaggerations. Hence, we could not detect less dramatic exaggerations that could account for some of the survey/UI earnings differences. For example, if a $5- an-hour job were reported as a $6-an-hour job, survey-reported earnings would be 20 percent higher than true earnings, but we could not detect this exaggeration.

[38] If we include all overtime earnings, the survey/UI earnings ratios are 1.30 for adult women, 1.40 for adult men, 1.43 for female youths, 1.62 for male youths with no prior arrest, and 1.93 for male youths with a prior arrest.

[39] We defined non-standard reporting periods to be other than hourly, daily, weekly, twice-a-week, or monthly. The survey asked respondents who were paid on an hourly basis to indicate their "regular hourly rate of pay", which we assume represented pre-tax wages. For job spells with non-hourly pay, respondents were asked whether they had reported it on a pre-tax or after-tax basis. Most reported pre-tax wages.

employment spells contained weeks not worked for pay because of illness or other reasons.[40] In the subsample of quarters with survey-reported earnings, the survey/UI earnings ratio was virtually the same for quarters that involved tips, bonuses, commissions or Aodd jobs@ and quarters that did not. In the same subsample, the survey/UI earnings ratio was almost the same for quarters with and without weeks not worked for pay.

The survey/UI earnings discrepancy was also unrelated to the industry of employment. These earnings ratios persistently exceeded one across manufacturing, wholesale and retail trade, services, or other industries. The higher earnings ratios for male youths could not be explained by the distribution of their jobs by industry.[41]

In contrast, job length and the presence of health insurance were related to the survey/UI earnings ratio. The earnings ratios for jobs lasting less than a year, and for those without health insurance, were somewhat higher than those for other types of jobs. This finding suggests that UI data may be more likely than surveys to miss earnings from short term, less substantial jobs.[42] But job length and insurance appears to explain only a small part of the overall survey/UI

---

[40] Less than 10 percent of the person-quarters with survey-earnings had any reported tips, bonuses or commissions, and less than 10 percent of total survey-earnings were from these sources. About 17-20 percent of the person-quarters with survey-earnings involved employment spells with weeks not worked, and the total number of time not worked was typically less than one month of the entire job spell. About one-fifth had wages from Aodd jobs@, although this comprised less than five percent of total survey-earnings.

[41] The industry of employment is only available for person-quarters involving the most recent job of a survey respondent. About 29 percent of these person-quarters involved earnings from manufacturing, 25 percent involved earnings from wholesale or retail trade, 35 percent involved earnings from service sector jobs, and 11 percent involved earnings from other industries.

[42] For person-quarters involving most-recent jobs with survey-reported earnings, the survey/UI ratio for jobs lasting less than a year versus those lasting longer was 1.43 versus 1.30 for adult women, 1.55 versus 1.37 for adult men, 1.68 versus 1.39 for female youth, 1.91 versus 1.58 for male youth without an arrest record, and 2.11 versus 2.05 for male youth with an arrest record. For jobs with health insurance versus those without it, the ratio was 1.45 versus 1.24 for adult women, 1.65 versus 1.25 for adult men, 1.68 versus 1.30 for female youth, 1.99 versus 1.42 for male youth without an arrest record, and 3.02 versus 1.58 for male youth with an arrest record

earnings discrepancy.[43]

VI. Our Findings in the Context of Previous Research

As described above, we found that data from quarterly employer reports to state UI agencies produced estimates of JTPA impacts on the earnings and employment of low-income persons that were comparable to those produced by data from retrospective surveys of individuals. Although average earnings reported by the survey data were somewhat higher than those reported by UI wage records, this difference was almost the same, on average, for program and control group members. Hence, it had virtually no affect on estimates of percentage earnings gains caused by the program, and it had only a modest affect on estimates of program-induced earnings gains measured in dollars. Furthermore, the two data-sources produced very similar measures of quarterly employment rates and estimates of JTPA impacts on employment rates..

These findings were based on the experience of a large and diverse sample of economically disadvantaged persons from 12 states and represent a number of policy-relevant groups, including: low-income adult men, low-income adult women, out-of school female youths, female AFDC recipients, adult UI claimants, and out-of-school male youths with no prior

[43] We can assess the importance of job length and health insurance coverage by first taking the sample of quarters with survey-reported earnings and increasing UI-reported earnings in person-quarters with short jobs (or without coverage) by a factor such that survey/UI earnings ratios in these quarters equal earnings ratios in quarters with long job spells (or with coverage). Next, we recompute overall earnings ratios for all quarters, given that the UI earnings in quarters with survey-reported short jobs (or coverage) are now higher. We then observe how much the overall, full-sample earnings ratios (from Exhibit 1) decline. If we increase UI-reported earnings in quarters with short survey-reported jobs in this fashion, earnings ratios for all person-quarters decline by a small amount -- to 1.19 for adult women, 1.25 for adult men and female youth, 1.41 for male youth without an arrest record, and 1.77 for male youth with an arrest record. If we increase UI-reported earnings in quarters with survey-reported health insurance in this manner, then earnings ratios for all person-quarters decline to 1.18 for adult women, 1.21 for adult men, 1.24 for female youth, 1.37 for male youth without an arrest record, and 1.46 for male youth with an arrest record.

arrests. The one striking exception to our general findings was for out-school male youths with prior arrests. Impact estimates from UI wage data and survey data were markedly different for this small subgroup.

The limited previous research which exists on this topic tends to support our central findings (see Exhibit 9). For example, in an unpublished presentation, Cave (1995) compared survey and UI estimates of impacts on quarterly earnings and employment rates for large samples of AFDC recipients from work/welfare demonstration programs in several states. He found that UI wage records and sample surveys produced similar impact estimates, although survey earnings were somewhat higher than UI data, on average. Cave also found that the two data-sources were more consistent in terms of quarterly employment rates than they were in terms of mean quarterly earnings.

Corson, et. al. (1991), Corson *et. al.* (1989), and Decker (1989) compared impact estimates from surveys and UI wage records for large samples of UI claimants from demonstration programs in New Jersey and Pennsylvania. They found that the two data-sources produced similar, although not identical, program impact estimates. In addition, they found that average survey earnings were slightly higher than UI earnings, but not by as much as we observed. The authors demonstrated that much of the survey/UI difference they found was due to out-of-state jobs and jobs not covered by UI.

Baj, Fahey and Trott (1993) report a detailed comparison of employment rates computed from survey data and UI data. Differences between these employment rates were generally small and comparable to those we observed. The authors conclude that UI wage records provide an

adequate substitute for surveys as the basis for monitoring ongoing job-training programs.

Lastly, Greenberg and Halsey (1983) compared estimates of the impact of a negative income tax on labor supply using survey data and UI data.[44] They found that impact estimates based on survey data were somewhat smaller than estimates based on UI data, although mean survey earnings for their sample were very close to mean UI earnings, especially for adults. They suggest that the difference between their survey-based impact estimates and their UI-based estimates reflected an incentive to under-report earnings that was correlated with the negative income tax rate for each sample member. This particular incentive does not exist in the context of most employment and training programs, however.

VII. Conclusions

This paper has compared estimates of program impacts on employment and earnings obtained from employer-reported UI wage records and from self-reported surveys of individuals. The analysis is based on a sample of 12,318 low income persons from 12 states for whom data from both sources were available. For adults and female youths, survey-reported average quarterly earnings were 23-36 percent higher than UI-reported average quarterly earnings. For male youths, the discrepancy was even larger. Survey- and UI-reported employment rates were

[44] There is an existing literature which is indirectly related to our paper. It examines the quality of survey-based earnings measures by comparing them mainly to employer records (not UI wage records) for a broad-based sample of workers (not just low-income persons). Specifically, Mellow and Sider (1983) compared earnings data from the January 1977 Current Population Survey with matched employer data; Duncan and Hill (1985) and Rodgers, Brown and Duncan (1993) compared employer-and employee-reported earnings for a single manufacturing firm; and Bound and Krueger (1991) compared 1977-78 earnings data from the Current Population Survey with earnings data from the Social Security Administration.

more similar.

For all adults, female youths, and the 76 perent of male youths with no prior arrest, the two data-sources provided estimates of JTPA impacts on earnings and employment that tell roughly the same story. The impact estimates from the two data-sources, when expressed as a percentage of the control group, are almost identical because the survey/UI earnings ratio is very similar for the program group and control group. The survey produces impact estimates that are larger in dollars because survey-reported average earnings are higher.

In contrast, for the 24 percent of male youth with a prior arrest record, the survey data indicate that JTPA markedly reduced average earnings, but UI data indicate the JTPA had no effect. This one glaring discrepancy reflects the fact that the survey/UI earnings ratio was much higher for control group members than it was for program group members.

We conducted a series of tests to explore the reasons for the discrepancies in reported earnings. Our findings persisted across all states. For all groups, the higher survey-reported average earnings are the net result of large positive and negative discrepancies in average earnings per person. Roughly half of the difference in average quarterly earnings appears to reflect earnings that are not reported to state UI agencies. Person-quarters with survey-reported jobs shorter than one year and with survey-reported medical insurance have smaller discrepancies in mean earnings. Much of the remaining discrepancy remains unexplained, however, and we could not explain the especially large discrepancies for male youths.

Our findings, plus those of previous researchers, suggest that UI wage records and individual follow-up surveys produce similar estimates of program impacts on the earnings and

employment **of the same individuals.** But UI wage records are much less expensive than sample surveys. To obtain a two-year (eight-quarter) earnings and employment history for an individual from a survey can cost over $100, whereas to obtain this same information from UI wage records can cost less than $1, and often only pennies.[45] Hence, UI wage records can be used to follow-up samples that are many times larger than those which could be followed-up by individual surveys.

A further advantage of UI wage records is that they are not subject to survey non-response, which can bias program impact estimates. Successful follow-up surveys for randomized experiments manage to interview about 75 percent to 85 percent of a study's sample. To the extent that labor market outcomes for survey respondents and non-respondents differ, and this difference is not the same for the program and control groups, survey-based impact estimates can be biased, even though random assignment was executed properly.[46]

Yet another advantage of UI wage records is their considerable potential for long-term follow-up. Because the costs of employment and training programs are incurred up-front, whereas their benefits (in terms of increased earnings and employment, or reduced welfare and UI receipt, *etc.*) can accrue over long periods of time, extensive follow-up is often required for a proper benefit-cost analysis.[47] Fortunately, once arrangements have been made to obtain UI wage

[45] We thank Irene Robling of the Manpower Demonstration Research Corporation for providing these estimates based on her extensive experience with both data-sources.

[46] In contrast, UI wage records are equally available for all sample members unless the program being tested affects the types of jobs taken by sample members in ways that differ for the program and control groups and are correlated with the probability of being reported to UI. Because the program group and control group survey/UI earnings ratio was very similar for all but one of the subgroups we examined, this "differential" UI coverage was quite rare.

[47] Unless short-term benefits are sufficient to offset program costs.

records, social security numbers have been obtained to identify these records for individual sample members, and software has been developed to extract the needed information, it is relatively simple to continue updating this information, perhaps indefinitely.

Nevertheless, although UI wage records have many advantages over surveys, the striking difference in findings from these two data-sources for male youths with prior arrests should raise a note of caution about their indiscriminant use. It seems plausible that youths in general, and youths with prior criminal records in particular,[48] might be especially likely to participate in informal labor market activities, work in the underground economy, and/or hold short-term jobs. Hence, UI wage records might be especially likely to miss some of their earnings. This could explain the very high survey/UI earnings ratio for male youths in general, and the fact that the survey/UI earnings ratio was much higher for male youths with a prior arrest than for other male youths. But it does not explain why the survey/UI ratio for male youths with prior arrests was mcuh higher for program group members than for control group members.

It is possible that access to JTPA programs "diverted" some program group members from their participation in jobs that are not reported by employers to state UI agencies but would be acknowledged by individuals during a survey. However, when we compared the detailed characteristics of jobs reported on the follow-up surveys by program and control group members (among male youths with prior arrests) there was no discernible difference in the nature or "quality" of their jobs.

---

[48] There were too few female youths with prior arrests to analyze this subgroup separately.

Because we cannot explain why the results for male youths with prior arrests were so different from those for all others in our sample, it is not clear how to anticipate subgroups which might present similar problems. Hence, it seems prudent to pursue a "mixed-mode" approach to future large-scale evaluations of employment and training programs. Researchers could rely on UI wage records as the primary follow-up data-source for their full sample plus individual surveys for a portion of the sample. Doing so would both facilitate a "cross-validation" of the UI wage records and provide more detailed information about labor market experiences (for a subsample) than would be possible using UI data only. Given the likely importance of UI wage records as a future data-source, additional research on the properties of these data is clearly warranted.

Exhibit 1

Estimates of Impacts on Mean Quarterly Earnings

Obtained from Survey Data and UI Data

	Mean Quarterly Earnings and Impacts		
	Survey	UI Data	Ratio
	(1)	(2)	(1)/(2)
Adult Women			
Program group mean	$1,294	$1,048	1.23***
Control group mean	1,141	922	1.24***
Impact in dollars	153***	126***	
Impact in percent[a]	13.4%	13.7%	
Adult Men			
Program group mean	1,917	1,456	1.32***
Control group mean	1,824	1,398	1.30***
Impact in dollars	94	59	
Impact in percent[a]	5.2%	4.2%	

Female Youths			
Program group mean	951	701	1.36***
Control group mean	949	700	1.36***
Impact in dollars	2	1	
Impact in percent[a]	0.0%	0.0%	
Male Youths without a prior arrest			
Program group mean	1,556	1,015	1.53***
Control group mean	1,655	1,103	1.50***
Impact in dollars	-99	-88	
Impact in percent[a]	- 6.0%	- 8.0%	
Male Youths with a prior arrest			
Program group mean	1,282	759	1.69***
Control group mean	1,531	760	2.01***
Impact in dollars	-248*	- 1	

Impact in percent[a]	- 16.2%	0.0%

[a] As a percent of the control group mean.

See exhibit notes on next page:

Exhibit 1 (cont'd)

Notes

The number of person-quarters (treatment, control) are: 18,275 and 8,916 for adult women; 13,329 and 6,482 for adult men; 9,452 and 4,316 for female youths; 5,009 and 2,442 for male youths without a prior arrest; 1,646 and 705 for male youths with a prior arrest.

The number of persons represented are: 4,943 adult women; 3,651 adult men; 2,113 female youths; 1,225 male youths without a prior arrest; 386 male youths with a prior arrest.

The test of whether a survey/UI ***ratio*** is significantly different from *one* is a t-test of whether the corresponding survey/UI ***difference*** in earnings is significantly different from *zero*. For this t-test, each person had a single observation of earnings, summed over available quarters. Each observation was weighted by the number of quarters for which the person was observed.

Impact estimates are the simple difference between average quarterly earnings of the treatment group and the control group. The test of the statistical significance of the impact is a simple difference of means t-test.

*** statistically significant at the 1 percent level; ** at the 5 percent level; * at the 10-percent level (for a two-tailed test.)

Exhibit 2

Estimates of Impacts on Quarterly Employment Rates

Obtained from Survey Data and UI Data

	Quarterly Employment Rate and Impacts, in Percentage Points		
	Survey (1)	UI Data (2)	Ratio (1)/(2)
Adult Women			
Program group rate	59.2%	57.6%	1.03***
Control group rate	54.5	54.1	1.01
Impact in percentage points	4.7***	3.5***	
Adult Men			
Program group rate	65.8	61.7	1.07***
Control group rate	63.5	60.7	1.05***

Impact in percentage points	2.4*	0.9	
Female Youths			
Program group rate	51.3	50.6	1.01
Control group rate	50.6	51.2	0.99
Impact in percentage points	0.6	-0.6	
Male Youths without a prior arrest			
Program group rate	65.5	61.3	1.07***
Control group rate	69.3	63.2	1.10***
Impact in percentage points	-3.8*	-1.9	
Male Youths with a prior arrest			
Program group rate	58.0	52.8	1.10**
Control group rate	61.3	55.0	1.11*
Impact in percentage points	-3.3	-2.2	

The samples and tests for this exhibit are the same as those for Exhibit 1.

*** statistically significant at the 1 percent level; ** at the 5 percent level; * at the 10-percent level (for a two-tailed test.)

Exhibit 3

Estimates of Impacts on Mean Quarterly Earnings Obtained from Survey Data and from UI Data: For Sample Members Who Were AFDC Recipients or UI Claimants When They Applied to JTPA

	Mean Quarterly Earnings and Impacts		
	Survey	**UI Data**	**Ratio**
	(1)	**(2)**	**(1)/(2)**
Adult Women, AFDC Recipients			
Treatment Group	$ 999	$ 789	1.27***
Control Group	791	636	1.24***
Impact in dollars	208***	153***	
Impact in percent[a]	26%	24%	
Female Youths, AFDC Recipients			
Treatment Group	716	535	1.34***
Control Group	670	531	1.26***

Impact in dollars	46	4	
Impact in percent[a]	7%	0%	
Adult Women, UI Claimants			
Treatment Group	$1,833	$1,597	1.15***
Control Group	1,861	1,515	1.23***
Impact in dollars	- 28	82	
Impact in percent[a]	1.5%	5%	
Adult Men, UI Claimants			
Treatment Group	2,488	1,959	1.27***
Control Group	2,281	1,920	1.19***
Impact in dollars	207	39	
Impact in percent[a]	9%	2%	

[a] As a percent of the control group mean.

The number of person-quarters for AFDC recipients (treatment, control) was: 6,679 and 3,415 for adult women, and 3,009 and 1,448 for female youths. The number of AFDC recipients represented was 1,796 adult women and 665 female youths. The number of person-quarters for UI Claimants was: 1,978 and 968 for adult women, and 2,257 and 1,206 for adult men. The number of UI claimants represented was 539 adult women and 648 adult men.

*** statistically significant at the 1 percent level; ** at the 5 percent level;

* at the 10-percent level (for a two-tailed test.)

Exhibit 4

Survey/UI Earnings Ratios by Site

JTPA Study Site	Adult Women	Adult Men	Female Youths	Male Youths without a prior arrest	Male Youths with a prior arrest
Corpus Christi, TX	**1.25**	1.43	1.30	**1.56**	**1.76**
Cedar Rapids, IA	1.25	1.25	1.43	1.40	1.73
Coosa Valley, GA	1.29	1.39	1.45	1.54	1.86
Heartland, FL	1.25	1.36	1.16	1.64	1.96
Fort Wayne, IN	1.15	1.24	1.24	1.38	1.58
Jackson, MS	1.34	**1.32**	1.31	1.61	3.28
Larimer County, CO	1.40	1.29	1.76	1.65	3.29
Decatur, IL	1.22	1.16	1.42	1.32	1.47
Northwest Minnesota	1.19	1.72	1.35	2.40	1.63
Omaha, NE	1.19	1.22	1.30	1.52	1.14

Providence, RI	1.21	1.31	**1.40**	1.57	1.62
Springfield, MO	1.34	1.49	1.71	1.47	2.96

The median ratio in each target group is shown in boldface.

Exhibit 5

Distribution of Mean Individual-Level Differences between Survey- and UI-Reported Quarterly Earnings

Mean Individual Survey-Reported Quarterly Earnings *minus* Mean Individual UI-Reported Quarterly Earnings	Adult Women	Adult Men	Female Youths	Male Youths without a prior arrest	Male Youths with a prior arrest
$2,001 or more					
$1,001 - $2,000	3.5%	9.9%	2.0%	8.7%	9.3%

$601 - $1,000	7.7	11.1	7.7	15.0	16.1
$401 - $600	7.9	9.4	9.1	12.4	13.2
$201 - $400	6.8	5.9	8.3	8.2	7.3
$1 - $200	10.4	8.6	13.0	11.9	9.6
$0	17.3	12.7	20.6	13.2	14.8
-$1 - -$200	14.3	8.1	10.0	3.8	5.7
-$201 - -$400	16.2	13.2	17.3	10.8	11.1
-$401 - -$600	6.2	6.2	5.3	5.6	6.0
-$601 - -$1,000	3.3	4.2	3.3	3.4	1.6
-$1,001 - -$2,000	3.4	4.6	2.0	4.5	3.4
-$2,001 or less	2.3	4.1	1.1	1.7	1.6
	0.9	2.0	0.2	0.8	0.5
Mean of Survey/UI difference per individual	$228	$451	$256	$547	$605
Survey/UI ratios					
Full Sample					

Without Outliers[a]	1.23	1.31	1.36	1.53	1.81
	1.20	1.29	1.33	1.51	1.80

[a] Without the largest 2.5 percent of positive differences and the largest 2.5 percent of negative differences.

The number of observations in this table is equal to the number of persons rather than person-quarters. Samples for each target group are the same as those for Exhibit 1.

Exhibit 6

Random and Systematic Components of Mean Individual-Level Differences between Survey-and UI-Reported Quarterly Earnings

	Adult Women	Adult Men	Female Youths	Male Youths without a prior arrest	Male Youths with a prior arrest
Root Mean Squared Difference	$875	$1,478	$709	$1,305	$1,237
Mean UI Earnings	$1,017	$1,432	$710	$1,034	$746
Percent of Root Mean Squared Difference Due To:					

Random Error					
Systematic Error	93.2	90.7	87.0	82.5	76.2
	6.8	9.3	13.0	17.5	23.8

The number of observations in this table is equal to the number of persons rather than person-quarters. Samples for each target group are the same as those for Exhibit 1

Exhibit 7

Mean Quarterly Earnings When Employment is Reported

by Both Data-Sources and by One Data Source Only

	One Source Reports Earnings			**Both Sources Report Earnings**		
	Survey Only (1)	UI Only (2)	Ratio (1)/(2)	Survey (4)	UI Data (5)	Ratio (4)/(5)
Adult Women	$1,825	$ 838	2.18***	$2,234	$1,973	1.13***
Adult Men	2,749	1,210	2.27***	2,941	2,579	1.14***
Female Youths	1,533	671	2.29***	1,956	1,582	1.24***
Male Youths without a prior arrest	2,224	961	2.31***	2,436	1,874	1.30***
Male Youths with a prior arrest						

Treatment Group						
Control Group	2,045	801	2.55***	2,292	1,665	1.37***
	2,574	723	3.56***	2,462	1,600	1.53***

Samples for each target group are the same as those for Exhibit 1.

BOTH SOURCES: The test of whether a survey/UI ***ratio*** is significantly different from one is a t-test of whether the corresponding ***difference*** in mean quarterly earnings is significantly different from zero. In this t-test, each person has a single observation of earnings summed over all available quarters; each person's single observation is weighted by the number of quarters for which he/she is observed.

ONE SOURCE: The test of the statistical significance of a survey/UI difference is a simple t-test of the difference between average earnings in the "survey-only" sample of person-quarters and average earnings in the "UI-only" sample of person-quarters. No weights are used.

*** statistically significant at the 1 percent level; ** at the 5 percent level; * at the 10-percent level (for a two-tailed test.)

Exhibit 8

Mean Quarterly Earnings from IRS Data and UI Data

	Mean Quarterly Earnings		
	IRS Data	UI Data	Ratio

	(1)	(2)	(1)/(2)
Adult Women	$1,295	$1,140	1.14***
Adult Men	1,897	1,541	1.23***
Female Youths	789	671	1.18***
Male Youths	1,061	847	1.25***

Samples for this exhibit are groups of persons for whom the IRS reported mean annual earnings. The exhibit only includes IRS groups that have valid annual UI earnings data for all members. Annual earnings are divided by four to obtain average quarterly earnings. The number of IRS groups included in the exhibit are: 261 for adult women; 253 for adult men; 196 for female youths; 148 for male youths.

The test of whether an IRS/UI ***ratio*** is significantly different from one is a t-test of whether the corresponding IRS/UI ***difference*** in mean quarterly earnings is significantly different from zero. In this t-test, each IRS group has a single observation.

*** statistically significant at the 1 percent level; ** at the 5 percent level; * at the 10-percent level (for a two-tailed test).

Exhibit 9.

Summary of Previous Related Studies

Exhibit 9.

Summary of Previous Related Studies

Study/Data Sources Used	Major Findings
Greenberg and Halsey (1983) SIME/DIME, 10th quarter of experiment; UI-based earnings adjusted upward slightly to reflect uncovered jobs	Survey/UI Mean Earnings Ratios, when both sources report earnings: Men: 1.07 ; Women: 1.05-1.14; Youth: 1.12-1.23 Ratios smaller for those subject to NIT Impact estimates: UI Data Yield Smaller Estimate of Work Reduction
Corson et al (1989),Decker (1989) Participants in New Jersey Unemployment Insurance Reemployment Demonstration, 4th quarter of experiment*	Survey/UI Mean Earnings Ratios: Full sample -- 1.14 After omitting self-employed -- 1.06 After omitting self-employed and those employed out-of-state -- 1.03 Estimates of impacts on earnings: Survey $201, UI $-33, neither statistically significant.
Corson et al (1991) Participants in Pennsylvania	Survey/UI Mean Earnings Ratios: Full sample -- 1.08 After omitting self-employed, those employed out-of-

Unemployment Insurance Reemployment Bonus Demonstration, 3rd quarter of experiment*	state, those with severance pay -- 1.05 Estimates of impacts on quarterly earnings: Slightly higher with survey data, but most estimates remain insignificant.
Baj, Fahey, and Trott (1993) 15-state survey of JTPA participants, UI data, and administrative JTPA data	Percent with Earnings Reported by UI Data Only: 11.3% Percent with Earnings Reported by Survey Data Only: 9.1%, of which: 15.3% employed out of state 6.6% uncovered jobs 23.1% delayed reporting or earn/pay lag
Cave, 1995 Summary of MDRC evaluation studies which used both survey and UI data	Survey/UI Mean Earnings Ratios: 1.05 -- 1.80 Survey/UI Ratios of Employment Rates: 0.8 -- 1.1 Estimates of impacts on earnings: Survey and UI yielded similar estimates.

* Severance pay reported to UI system as earnings made UI-based average earnings relatively higher in 1st quarter of experiment.

REFERENCES

Baj, John, Sean Fahey, and Charles Trott. 1992. "A Study of the Use of Unemployment Insurance Wage-Record Data as an Evaluation Tool for JTPA: Phase II Activities." ***Using Unemployment Insurance Wage-Record Data for JTPA Performance Management.*** Washington, DC: National Commission for Employment Policy.

Baj, John, Charles Trott, and David Stephens. 1991. ***A Feasibility Study of the Use of Unemployment Insurance Wage-Record Data as an Evaluation Tool for JTPA: Report on Project's Phase I Activities.*** Washington, DC: National Commission for Employment Policy.

Bell, Stephen, and Larry L. Orr. 1994. "Is Subsidized Employment Cost Effective for Welfare Recipients? Experimental Evidence from Seven State Demonstrations." ***Journal of Human Resources 29*** (Winter): 42-61.

Betsey, Charles L., Robinson G. Hollister, and Mary R. Papageorgiou. 1985. ***Youth Employment and Training Programs: The YEDPA Years.*** Committee on Youth Employment Programs, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press.

Bloom, Howard S. 1990. ***Back to Work: Testing Reemployment Services for Displaced Workers*.** Kalamazoo, MI: W.E. Upjohn Institute.

Bloom, Howard S., Larry L. Orr, Stephen H. Bell, George Cave, Fred Doolittle, Winston Lin, and Hans Bos (forthcoming). "The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National JTPA Study." ***The Journal of Human Resources*.**

Bloom, Howard S., Larry L. Orr, George Cave, Stephen H. Bell, and Fred Doolittle. 1993. ***The National JTPA Study: Title II-A Impacts on Earnings and Employment at 18 Months*.** Bethesda, Md.: Abt Associates Inc.

Bound, John, Charles Brown, Greg Duncan, and Willard Rodgers. 1990. "Measurement Error in Cross-Sectional and Longitudinal Labor Market Surveys: Validation Study Evidence." ***Panel Data and Labor Market Studies***, eds. J. Hartog, G. Ridder, and J. Theeuwes. Amsterdam: North Holland.

Bound, John, and Alan Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" ***Journal of Labor Economics*** *9* (January): 1-24.

Cave, George, MDRC. 1995. "Partially Overlapping Data Sources: Data Plans for the Parents' Fair Share Evaluation." Presentation, APPAM Conference, November, Washington, DC.

REFERENCES

(Continued)

Corson, Walter, Paul Decker, Shari Dunstan, and Anne Gordon. 1989. ***The New Jersey Unemployment Insurance Reemployment Demonstration Project: Final Evaluation Report.*** Princeton, NJ: Mathematica Policy Research, Inc.

Corson, Walter, Paul Decker, Shari Dunstan, and Stuart Kerachsky. 1991. ***Pennsylvania Reemployment Bonus Demonstration: Final Report.*** Princeton, NJ: Mathematica Policy Research, Inc.

Decker, Paul. 1989. "Systematic Bias in Earnings Data Derived from Unemployment Insurance Wage Records and Implications for Evaluating the Impact of Unemployment Insurance Policy on Earnings." Princeton, NJ: Mathematica Policy Research.

Duncan, Greg J., and Daniel Hill. 1985. "An Investigation of the Extent and Consequences of Measurement Error in Labor-economic Survey Data." ***Journal of Labor Economics*** *3* (October): 508-532.

Edin, Kathryn. 1993. ***There's a Lot of Money Left at the End of the Month: How Welfare Recipients Make Ends Meet in Chicago***. New York: Garland.

Fein, David J., Erik Beecroft, and John D. Blomquist. 1994. ***The Ohio Transitions to Independence Demonstration: Final Impacts for JOBS and Work Choice.*** Bethesda, Md.: Abt Associates Inc.

Greenberg, David, and Harlan Halsey. 1983. "Systematic Misreporting and Effects of Income Maintenance Experiments on Work Effort: Evidence from the Seattle-Denver Experiment." ***Journal of Labor Economics*** *1* (October): 380-407.

Gueron, Judith M., and Edward Pauly, 1991. ***From Welfare to Work.*** New York: Russell Sage Foundation.

Hamilton, William L., Nancy R. Burstein, Margaret Hargreaves, David A. Moss, and Michael Walker. 1993. ***The New York State Child Assistance Program: Program Impacts, Costs, and Benefits.*** Cambridge, Mass.: Abt Associates Inc.

Kemple, James J., and Joshua Haimson. 1994. ***Florida's Project Independence: Program Implementation, Participation Patterns, and First-Year Impacts.*** New York: Manpower

Demonstration Research Corporation.

Mellow, Wesley, and Hal Sider. 1983. "Accuracy and Response in Labor Market Surveys: Evidence and Implications." ***Journal of Labor Economics*** *1* (October): 331-343.

REFERENCES

(Continued)

Meyer, Bruce. 1995. "Lessons from the U.S. Unemployment Insurance Experiments." ***Journal of Economic Literature.*** *1* (March): 91-131.

Orr, Larry L., Howard S. Bloom, Stephen H. Bell, Fred Doolittle, Winston Lin, and George Cave. 1996. ***Does Training for the Disadvantaged Work? Evidence from The National JTPA Study*****.** Washington, DC: The Urban Institute Press.

Orr, Larry L., Howard S. Bloom, Stephen H. Bell, Winston Lin, George Cave, and Fred Doolittle. 1994. ***The National JTPA Study: Impacts, Benefits, and Costs of Title II-A*****.** Bethesda, Md.: Abt Associates Inc.

Puma, Michael, and Nancy Burstein. 1994. "The National Evaluation of the Food Stamp Employment and Training Program." ***Journal of Policy Analysis and Management*** *13* (Spring): 311-331.

Riccio, James, Daniel Friedlander, and Stephen Freedman. 1994. ***GAIN: Benefits, Costs, and Three-Year Impacts of a Welfare-to-Work Program.*** New York: Manpower Demonstration Research Corporation.

Rodgers, Willard, Charles Brown, and Greg Duncan. 1993. "Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages." ***Journal of the American Statistical Association*** *88* (December): 1208-1218.

Spiegelman, Robert G., Christopher J. O'Leary, and Kenneth J. Kline. 1992. ***The Washington Reemployment Bonus Experiment: Final Report.*** Unemployment Insurance Occasional Paper 92-6. Washington, DC.: U.S. Department of Labor, Employment and Training Administration.

Stromsdorfer, E., H. Bloom, R. Boruch, M. Borus, J. Gueron, A. Gustman, P. Rossi, F. Scheuren, M. Smith, and F. Stafford. 1985. ***Recommendations of the Job Training Longitudinal Survey Research Advisory Panel.*** Washington, D.C.: Employment and Training Administration, U.S. Department of Labor.

U.S. Congress, Office of Technology Assessment. May 1994. ***Wage Record Information Systems.*** Washington, DC: OTA-BP-EHR-127.

Werner, Alan, and Robert Kornfeld. 1996. ***The Evaluation of To Strengthen Michigan Families: Third-Year Impacts.*** Bethesda, Md.: Abt Associates Inc.