EXECUTIVE SUMMARY

This evaluation is motivated by two goals: (1) to assess the reliability of the impact estimates provided in the evaluation of the Worker Profiling and Reemployment Services (WPRS) programs, and (2) to compute revised estimates of the impacts of WPRS programs if a more accurate estimation method can be identified. The evaluation also provides general information on the accuracy of different methods for estimating impacts without random assignment.

Under WPRS, states were required to establish systems for identifying Unemployment Insurance (UI) claimants likely to exhaust their UI benefits and referring them to reemployment services, such as resume preparation and training in job search methods. In an evaluation sponsored by the U.S. Department of Labor (USDOL), the impacts of WPRS were estimated by comparing UI claimants who were assigned to WPRS services (the treatment group) to claimants who were not assigned to WPRS services (the comparison group). Linear regression techniques were used to control for pre-existing differences between the two groups.

The results from the WPRS evaluation suggest that the impacts of WPRS on earnings are positive in some states and negative in others. However, the wide variation in impact estimates across states raises questions about the accuracy of the estimates. Furthermore, when the pre-existing differences between the treatment and comparison groups are large--as in the WPRS evaluation--linear regression methods can be unreliable. Therefore, the wide state-to-state variation in the estimated earnings impacts *may* be due to estimation error attributable to the regression method used in the WPRS evaluation.

Prior to the implementation of WPRS, USDOL sponsored a demonstration to test different program models that are consistent with the regulations governing WPRS. In 1995, the Job Search Assistance (JSA) Demonstration was implemented in the District of Columbia and in selected counties in Florida. Because the demonstration was based on the random assignment of eligible claimants to treatment and control groups, impacts were estimated by comparing treatment group members to control group members. Random assignment ensured that the pre-existing differences between the two groups were negligible.

Therefore, the demonstration should provide reliable estimates of the impacts of different WPRS program models via treatment-control differences. Furthermore, demonstration data can be used to compute other impact estimates using data that mimic the treatment and comparison samples available to the WPRS evaluation. The reliability of these impact estimates can be tested by comparing them to the treatment-control differences.

In this evaluation, we use data from the JSA Demonstration in Florida to mimic the treatment and comparison samples from the WPRS evaluation, and to test different methods of estimating impacts from these samples. These methods include the regression method used in the WPRS evaluation, but also include variants of the matching methods used in other evaluations. Matching is designed to select a subgroup of comparison group members who are similar to treatment group members. Impacts are then estimated by comparing treatment group members to the subgroup of similar comparison group members.

The plan for the evaluation included two phases:

- ***Phase I: Testing Different Methods of Estimating Impacts Using JSA Data.*** In Phase I, use data from the JSA Demonstration to assess the reliability of the regression method employed in the WPRS evaluation and the matching methods developed in this evaluation.
- ***Phase II: Applying Matching Methods to Actual WPRS Data.*** If any of the matching methods produce more accurate impact estimates than the regression method, apply those matching methods to WPRS data to obtain revised estimates of the impacts of WPRS on earnings.

DESIGN OF PHASE I OF THE EVALUATION

The design of Phase I consisted of two components: (1) identifying the analysis samples from JSA Demonstration data; and (2) specifying methods for estimating the impacts of being assigned to JSA/WPRS services on the claimants *who would have been assigned to services if WPRS had been operating in Florida in place of the demonstration*.

Identifying Three Samples from JSA Demonstration Data. We used the rule by which UI claimants are assigned to WPRS to determine which claimants would have been assigned to WPRS had it been operating instead of the demonstration. Claimants who would have been assigned to WPRS were classified as "treatment claimants" or "control claimants" for this evaluation based on their treatment-control status in the demonstration. Claimants who would *not* have been assigned to WPRS (and were not treated in the demonstration) were classified as "comparison claimants".

Specifying the Methods for Estimating Impacts. Based on the three analysis samples, we specified alternative methods of estimating the impacts of being assigned to WPRS. The experimental benchmark estimate equals the mean earnings of treatment claimants minus the mean earnings of control claimants. This benchmark is used to assess whether accurate impact estimates can be computed from "nonexperimental data"--data on treatment and *comparison* claimants--using either the linear regression method from the WPRS evaluation or one of the matched comparison groups developed for this evaluation.

The matching methods developed for this evaluation are designed to select "matched comparison groups" that look like the treatment group. A comparison claimant is selected for the matched comparison group if he or she can be "matched" to one or more treatment claimants with similar characteristics. The rules developed for defining acceptable matches require that matched claimants have the same sex, race/ethnicity, and education. Furthermore, matching claimants must have similar values for one of the following three variables:

1. ***Profiling Score.*** UI claimants are assigned "profiling scores" that reflect the probability of exhausting UI benefits without additional reemployment services, and are assigned to WPRS based on these scores.

2. ***Base-year Earnings.*** Claimants are determined eligible for UI based on their "base-year earnings", which measures total earnings in four out of five quarters prior to the UI claim.

3. ***Propensity Score.*** Treatment claimants have higher probabilities or propensities of being assigned to services than comparison claimants, and "propensity scores" are often computed in evaluations to use as matching variables.

FINDINGS FROM PHASE I OF THE EVALUATION

Based on the treatment and control groups in this evaluation, the experimental benchmark estimate that we use to assess the accuracy of other impact estimates equals $260. Therefore, the average earnings of treatment claimants in the year following the quarter of random assignment were $260 higher than the average earnings of control claimants in the same year.

How well did the different methods for estimating earnings impacts from the treatment and comparison samples perform? The two main findings from Phase I of the evaluation are given below:

1. ***The linear regression model used in the WPRS evaluation produced accurate impact estimates.*** The estimate produced by the linear regression model from the WPRS evaluation equals $308, which is very close to the experimental benchmark of $260.

2. ***The matched comparison groups tested in this evaluation produced less accurate impact estimates than the linear regression model.*** The impact estimates based on matched comparison groups range from -$111 to -$3,440, and none of these estimates are as close to the experimental benchmark as the estimate produced by the linear regression model.

Therefore, despite the general concerns that can be raised about the reliability of regression methods to adjust for large differences between treatment and comparison groups, *this evaluation provides no evidence that the regression methods used in the WPRS evaluation are unreliable*.

The poor performance of the matching methods tested in this evaluation can be attributed to the difficulty in selecting matched comparison groups that are sufficiently similar to the treatment group. Each matched comparison group was similar to the treatment group on many dimensions but different from the treatment group in at least one dimension that proved to be important. None of the matched comparison groups had the same (or a very similar) distribution of claimants across the local offices in the demonstration as the treatment group. Findings in this report suggest that it may be impossible to create a matched comparison group that is comparable to the treatment group in the distribution of claimants across local offices, and is *also* comparable to the treatment group in other important dimensions, such as sex, race/ethnicity, education, the profiling score, base-year earnings, and the propensity score. In other words, we were unable to create a matched comparison group that was comparable to the treatment group on all the dimensions that seemed important.

Based on the results from Phase I, MPR recommended not to proceed to Phase II of this evaluation, and USDOL concurred. This recommendation was based on the finding that the regression methods used in the WPRS evaluation produced accurate estimates of earnings impacts from the demonstration data. This report provides no evidence that matched comparison groups of the types we tested would yield more accurate estimates of the impacts of WPRS.