

<b>Employment and Training Administration</b> <b>Advisory System</b> <b>U.S. Department of Labor</b> <b>Washington, D.C. 20210</b>	<b>CLASSIFICATION</b> UI
	<b>CORRESPONDENCE SYMBOL</b> OWS/DFAS
	<b>DATE</b> January 6, 2006

**ADVISORY:** UNEMPLOYMENT INSURANCE PROGRAM LETTER NO. 9-06

**TO:** STATE WORKFORCE AGENCIES

**FROM:** CHERYL ATKINSON *Cheryl Atkinson*  
 Administrator  
 Office of Workforce Security

**SUBJECT:** Survey and Data Collection for the Evaluation of Worker Profiling and Reemployment Services (WPRS) Models

1. Purpose. To inform state workforce agencies (SWAs) that the Office of Management and Budget (OMB) has approved a one-time collection of information for the evaluation of worker profiling models, to provide data reporting formats, and to provide reporting instructions.
2. References. Unemployment Insurance Program Letter No. 13-94, Unemployment Insurance Program Letter No. 41-94, and Federal Register Notice, Vol. 70, No. 102, pages 30815-30816.
3. Background. The Unemployment Insurance (UI) WPRS program was established to identify, as early as possible in the claims series, those beneficiaries who are likely to exhaust their benefits in order to provide them services helpful in obtaining employment. The identification of these beneficiaries may be accomplished through the use of a statistical model or a characteristic screen. To date, no comprehensive analysis of the predictions made by these models has been undertaken. The information we are requesting in this data collection will be used for a study, the goals of which are to improve state worker profiling models by: 1) establishing an approach for evaluating the accuracy of worker profiling models; 2) applying this approach to current state models to determine how effective they are at predicting UI benefit exhaustion; and 3) based on the results, developing guidance on best practices in operating and maintaining worker profiling models.
4. OMB Approval. This data collection has been approved by OMB under the Paperwork Reduction Act of 1995, OMB approval number 1205-0454, expiration date: 11/30/2008. SWAs are not required to respond to this collection of information unless it displays a valid OMB control number.

<b>RESCISSIONS</b> None	<b>EXPIRATION DATE</b> January 6, 2007
----------------------------	---

5. Disclosure Statement. Based on information obtained from SWAs, it will take an estimated 32 hours to complete the survey and compile the data requested. Obligation to furnish this information is mandatory under Sec. 303(a)(6) of the Social Security Act.
6. Data Required. Attachment A, titled *Data Requirements for the Evaluation of State Worker Profiling Models*, contains instructions for selecting and preparing the content and format of the data required for analyzing the performance of states' worker profiling models and screens. Attachment B, titled *Worker Profiling and Reemployment Services Survey*, contains questions about operational and structural features of the worker profiling model or screen.
7. Action Required. Please submit the required data set and completed survey by email to the attention of Scott Gibbons at the address shown below. The data and completed survey are due no later than February 15, 2006. If electronic transmission by email of the survey and data are not possible, please contact Scott Gibbons for alternative reporting instructions.
8. Inquiries. All questions or inquiries should be directed to Scott Gibbons at [gibbons.scott@dol.gov](mailto:gibbons.scott@dol.gov) or 202-693-3008 (this is not a toll free number).

Attachments

Attachment A: Data Requirements for the Evaluation of State Worker Profiling Models

Attachment B: Worker Profiling and Reemployment Services Survey

# DATA REQUIREMENTS FOR THE EVALUATION OF STATE WORKER PROFILING MODELS

## Introduction

The goal of this project is to measure and analyze the accuracy of state Worker Profiling and Reemployment Services (WPRS) models. Summarized claimant data will be used to identify, to the greatest extent possible given the available data and analysis results, factors correlated with good or poor model performance.

## Required Data

Since different WPRS models and systems are found throughout the country, the specific data elements submitted will be slightly different for each state. This study requires three general types of information for each claimant. The first is the set of **data elements used to predict benefit exhaustion** at the time of the initial claim. These will be used to replicate the probability assigned to the claimant at the time of the initial claim and form the basis of the predicted outcome. The second general type of data is a **measure of benefit receipt** that will form the basis for the actual outcome. The last general type of data is an **indication of whether the claimant was referred to services** under WPRS. The analysis will compare the actual and predicted outcomes and explore the reasons for different levels of accuracy.

## Data Elements Used To Predict Benefit Exhaustion

The data elements provided for this study should be those used in the state WPRS model to compute probabilities of exhaustion. For example, if your state model uses industry, education and wage replacement rate, you would provide these data elements for all claimants with benefit year begin dates in the specified time range. For variables requiring categorization, you may either provide the value before categorization or the binary categorical variables, whichever is easier. For computed variables such as wage replacement rate, the state may provide the computed variable or the data elements used to compute it and the computation method. Data for the variables in the model and computation methods should be consistent with the information provided on the WPRS Survey in questions 22 and 24.

Collection of these data elements is necessary because a critical part of the subsequent analysis will correlate the predictive variables or categories with model performance. However, if states can additionally provide the predicted value assigned by the model or screen at the time of the initial claim, this would be very helpful in validation of the replicated state model.

## **Data On Benefit Receipt By Claimants**

In order to help determine whether the initial prediction of benefit exhaustion was a good one, a measure of benefit utilization is necessary. There are several ways that the state could provide this information, but the preferred approach is to provide the maximum potential benefit amount and the amount of benefits paid over the benefit year for each claimant.

## **Data On Referral To Services**

The final data set should include a single variable that indicates whether each claimant was referred to mandatory employment services through the WPRS program. This should be a simple indicator variable and can be coded as “yes/no”, “0/1” or any other suitable system.

## **Selecting Claims And Claimants**

Some states profile all claims, while other states exclude certain claimants (e.g., job attached) from worker profiling entirely. The data provided for this project should contain information for only those claimants who would be assigned scores by the profiling model and be eligible for referral under WPRS. For example, if the state does not profile additional or transitional claims, then those data should be excluded. Similarly, if the state does not normally create profiling scores for job attached claimants, data for those individuals should be excluded as well. In general, the subset of claims requested should be in agreement with answers to questions 11 and 26 on the WPRS Survey.

States are requested to provide these data for all claimants who meet the above criteria and whose benefit years began in the period 01/01/2003 through 12/31/2003. To provide an analysis of a period with more robust growth, data are also requested for claimants who meet the above criteria and whose benefit years began in the period 01/01/1999 to 12/31/1999. If claimant data for 1999 is not readily available due to Y2K computer changes or archival processes, claimant data for 2000 may be substituted. Please submit data for each year as a separate file.

## **Data Format**

One record should be submitted for each claimant who meets the criteria described above. The record should include the variables used to compute profiling scores, the maximum benefit amount and the total amount of benefits paid over the course of the benefit year. The data should be submitted in ASCII format with a brief description of the field names and widths. Please note that no personal identifiers should be present in the data that are submitted.

Below is an example of the data file from a hypothetical state that uses wage replacement rate, industry categorized, and years of education categorized to predict benefit exhaustion.

Case	Maximum Benefit Amount	Benefits Drawn	WPRS score	Referred To Services	Years Of Educ.	NAICS code	Weekly Benefit Amount	Base Period Wage
1	6500	5000	0.56	Y	16	623546	250	29000
2	5850	5850	0.86	Y	12	243567	225	23400
3	5980	2760	0.36	N	15	734539	230	28000

In this example, the state has elected to provide the raw data for industry (NAICS code) and education (years of education) as opposed to the categorized forms. The state would have provided information in its survey response on how these variables are categorized; therefore, the categorical forms of the variables could easily be re-created.

The state has also chosen to provide the raw data for wage replacement rate rather than the computed variable. The survey response would show how the state computes each claimant’s wage replacement rate. In this example, the state computes wage replacement rate as weekly benefit amount divided by the base period wages divided by 52; therefore, the state has provided both of these variables.

To help determine the actual claimant outcomes, data are also provided for the maximum benefit amount in dollars and the amount of benefits paid to the claimant in dollars. The estimated probability (score) that the WPRS model calculated for each claimant is included to help verify that the re-created model is consistent with the state calculations. Similarly, a single variable, coded as “yes” or “no” is included that indicates whether each claimant was referred to services.

If the state elects to provide the categorized and computed variables, the data file might look like this:

Case	Maximum Benefit Amount	Benefits Drawn	WPRS score	Referred To Services	Categorized Education	Industry Category	Wage Replacement Rate
1	6500	5000	0.56	Y	4	6	0.448
2	5850	5850	0.86	Y	2	2	0.500
3	5980	2760	0.36	N	3	7	0.427

The survey asks detailed questions regarding how these data are converted into probabilities; therefore sufficient information should be present to allow an accurate replication of the state model.

If you have questions regarding this data request, please contact Scott Gibbons at [gibbons.scott@dol.gov](mailto:gibbons.scott@dol.gov) or 202-693-3008.

# Worker Profiling Reemployment Services (WPRS) Survey

## *(Operational Section)*

Please enter the name of your State: \_\_\_\_\_

1. Please provide the name, title, e-mail address, and phone number of the individual(s) completing this survey including which survey questions they completed:

---

---

---

---

2. Please provide the name, title, e-mail address, and phone numbers of the individuals within UI, ES (Workforce Development), LMI, and IT who provide daily control and oversight of the WPRS process and model (if different from above).

---

---

---

---

3. How frequently is the model updated (run to generate new statistical parameters)?

Yearly \_\_\_\_\_  
2-3 Years \_\_\_\_\_  
More than 3 Years \_\_\_\_\_  
Other \_\_\_\_\_

3a. Date of Last Update: \_\_\_\_\_

4. Has the model been revised (i.e., **other than update**, has the model been revised in any way, such as a change in the variables used, the variable definitions, or functional form) since implementation?

Yes \_\_\_\_\_  
No \_\_\_\_\_

4a. If Yes, please provide date of last revision and brief description of revisions made:

---

---

---

---

---

4b. Do you have policy guidance to revise your model and if so, how often? Is there a decision maker within your agency who determines that the model will be revised?

---

---

---

---

5. By which method(s) is your initial claim process performed? (check all that apply and estimate percentages)

In-Person \_\_\_\_\_  
By Telephone \_\_\_\_\_  
By Mail \_\_\_\_\_  
Internet \_\_\_\_\_  
Other: (specify) \_\_\_\_\_

---

6. Are all of the claimant "characteristics" data needed for profiling purposes captured at the time of the initial claim?

Yes \_\_\_\_\_  
No \_\_\_\_\_

6a. If you answered "No" above, please describe how, and when, the data are captured or generated:

---

---

---

6b. Are there any checks on the accuracy of claimant provided information?

Yes \_\_\_\_\_ (If Yes, Please describe below)  
No \_\_\_\_\_

---

---

---

7. How frequently is the WPRS model run?

Daily \_\_\_\_\_

Weekly \_\_\_\_\_

Other (please describe) \_\_\_\_\_

\_\_\_\_\_

7a. Is the listing of profiling candidates produced at the same time the model is run?

Yes \_\_\_\_\_

No \_\_\_\_\_

If "No," please describe when the listing is produced:

\_\_\_\_\_

8. Is the model run against the first pay records?

Yes \_\_\_\_\_

No \_\_\_\_\_

8a. If you answered "No" to question 8, please describe against what UI or other data the model is run.

\_\_\_\_\_

9. Who determines and assigns the claimant's occupational code?

Initial Claims Taker \_\_\_\_\_

Workforce Dev. Worker \_\_\_\_\_

Other (please describe) \_\_\_\_\_

\_\_\_\_\_

9a. Which occupational coding system is used (DOT, SOC or, if any other classification system is used, please identify)?:

\_\_\_\_\_

9b. How is the occupational code derived for the claimant? (please describe, if not a standard classification system)

\_\_\_\_\_

\_\_\_\_\_



10. How is the claimant's primary employer (for assigning NAICS/SIC code) determined?

Review of work history with claimant \_\_\_\_\_

Review of wage records \_\_\_\_\_

Other (please describe) \_\_\_\_\_

11. Who is exempt from profiling in your State?

---

---

---

12. To whom is the list of profiling candidates sent, and using what medium? (describe)

---

---

13. Who determines the number of profiled candidates to be served and how is the number determined?

---

---

---

14. How do the probability scores, or rankings, influence selection of candidates from the pool?

---

---

---

15. Under what conditions can the local area skip down ranks in selecting candidates for services?

---

---

---

16. Are there feedback loops in place between local area operations and the WPRS model builders?

Yes \_\_\_\_\_

No \_\_\_\_\_

17. The original parameters for WPRS suggested individuals who had received more than 5 weeks of benefits prior to selection be excluded from the pool (e.g. if payment delays have deferred first payments for more than five weeks). Is this parameter in place in your system?

Yes \_\_\_\_\_  
No \_\_\_\_\_ If No, please provide the number you use \_\_\_\_\_

18. Has the accuracy of data needed for the Characteristic Screens been measured or tested to compare it to the predictive equation approach or has the existence of missing or inaccurate data been investigated?

Yes \_\_\_\_\_ (If Yes, please describe results below)  
No \_\_\_\_\_

---

---

---

19. Has your agency conducted any studies to evaluate the accuracy of the profiling model in predicting who will exhaust benefits?

Yes \_\_\_\_\_ (If Yes, please describe below)  
No \_\_\_\_\_

---

---

---

**(Structural Section)**

**SOME STATES MAY FIND IT ADVANTAGEOUS TO SIMPLY ATTACH TECHNICAL REPORTS OR COMPUTER PRINT-OUTS TO REPLY TO THE HIGHLY TECHNICAL STRUCTURAL QUESTIONS (especially 24, 25, 26, & 31). PLEASE BE SURE TO ATTACH THE REPORTS AND EXPLAIN WHERE IN THE REPORT OR PRINTOUT THE PERTINENT MATERIAL MAY BE FOUND.**

**Please note all questions that follow apply to the model that was primarily in use during the period \_\_\_\_\_ to \_\_\_\_\_.**

20. Which type of WPRS Model does your state currently use? Enter "Yes" in appropriate block.

Characteristic Screen \_\_\_\_\_  
Statistical Model \_\_\_\_\_

20a. What is your model's functional form? (example: logit, probit, tobit, linear, characteristic screen, other). \_\_\_\_\_  
\_\_\_\_\_

21. Which individuals are included in the data when the model was first estimated, or when it was updated or revised?

All initial claim filers \_\_\_\_\_  
Only benefit recipients \_\_\_\_\_  
Union member \_\_\_\_\_  
Others not profiled (describe) \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

21a. What is the sample size in the model's latest update and what was the original sample size when the model was first estimated?  
\_\_\_\_\_  
\_\_\_\_\_

22. What is your model's dependent (left-hand side) variable?

Exhaustion \_\_\_\_\_  
Duration of Benefits \_\_\_\_\_  
Both \_\_\_\_\_  
Other (describe) \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

23. For the purpose of updating your model, how do you define exhaustion of benefits? (check all options which apply)

Maximum benefits paid \_\_\_\_\_  
Received 26 weekly payments \_\_\_\_\_  
Benefit payments denied but under appeal \_\_\_\_\_  
Other (describe) \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

24. What are your model's independent (right-hand side) variables and how are they defined? Please include and explain how to calculate the variables and explain what data are used to create the variable. (examples: maximum duration = maximum benefit amount divided by weekly benefit amount; example 2: industry = the first digit of the NAICS hierarchical code). If you use a characteristic screen, what characteristics do you use?

---

---

---

---

25. What determined the selection of the independent (right-hand side) variables used in the predictive equation? Were any other variables examined?

---

---

---

---

25a. What are the numerical values of the estimated coefficients for the independent (right-hand side) variables and if this information is readily available, what was the standard error for each? (This information should be found on the original statistical output for the original estimation technique.)

---

---

---

---

26. What techniques are followed to prepare a claims data record for profiling using the equation? (e.g., are there checks for missing values, are union member claims profiled, are there data quality checks, etc.)

---

---

---

---

27. How are claimants with incomplete records, or records with missing variables, processed? (check all that apply)

- a. variable kept blank and a binary variable used to track the missing variable
- b. another version of the profiling model used
- c. value of missing data estimated by some other procedure
- d. missing value replaced by average value for the individuals in the run or some other average value
- e. Other method? (please describe) \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

28. Were the exclusion rules (see question 11) applied to the data records used during the estimation of the predictive equation? That is, were records excluded from the estimation database, and what percentage of claimants is excluded from profiling?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

29. Were the data quality procedures that were used for the data in the estimation of the predictive equation different from those used now for profiling? If so, how? In your view, does the elimination of claim records as a result of data quality procedures have an effect (either negative or positive) on the performance of the equation?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

30. Are the predicted values of the dependent (left-hand side) variables retained in electronic storage archives?

Yes \_\_\_\_\_  
No \_\_\_\_\_

31. What are the ranges of permissible (or expected) values of the data for the independent (right-hand side) variables (minimum and maximum)? Please describe below.

---

---

---

---

---

---

TEUFA:SGibbons:ajc:01/06/05:C4526:693-  
3008:S:\OIS\DFAS\ACTS\SCOTT\profile\wids\PBB\Draft WPRS Survey.doc  
Cc: ETACS/EOC/ OCIA/Atkinson/ Wilus/ Miller