12-2002

Comments on: Procedure for Profiling Unemployed Citizens in Russia

Christopher J. O'Leary
W.E. Upjohn Institute for Employment Research, oleary@upjohn.org

Comments on: Procedure for Profiling Unemployed Citizens in Russia

Authors
Christopher J. O'Leary, *W.E. Upjohn Institute for Employment Research*

Upjohn Author(s) ORCID Identifier
https://orcid.org/0000-0002-3372-7527

Comments on:

Procedure for Profiling Unemployed Citizens in Russia

December 2002

prepared for

Centre for Economic and Financial Research (CEFIR)
Nakhimovsky prospect 47
Office # 720
Moscow 117418
Russian Federation
Tel: +7 (095) 105 5002
Fax: +7 (095) 105 5003
cefir@cefir.ru
www.cefir.org

prepared by

Christopher J. O'Leary
W. E. Upjohn Institute for Employment Research
300 South Westnedge Avenue
Kalamazoo, Michigan 49007, USA
Tel: (269) 343-5541
Fax: (269) 343-3308
oleary@upjohn.org
www.upjohn.org

O'Leary is senior economist at the W.E. Upjohn Institute for Employment Research. Opinions expressed in this report are solely those of the author and do not reflect opinions of either the Upjohn Institute or its board of trustees.

Background for Comment on the Proposed Russian Profiling System

Uses of profiling in the United States

There are two main types of profiling for employment services used in the United States (US) today. The first application was the Worker Profiling and Reemployment Services (WPRS) system implemented following 1993 federal legislation (Wandner 1994). WPRS requires all states to have a methodology for early identification of unemployment insurance (UI) beneficiaries at risk of long term unemployment and to refer those with the highest risk to early reemployment services (Dickinson et al. 1999, 2002). Several evaluations of work search requirements and job search assistance (JSA) for UI beneficiaries have found low cost of services and high cost effectiveness (O'Leary 2002). The WPRS system simply refers those most likely to exhaust UI to a special set of JSA, with benefits suspended for those invited and failing to attend. The WPRS models are also used as part of the referral process for self-employment assistance under UI (Messenger et al. 2002).

The Frontline Decision Support System (FDSS) is a pilot system being tested in one-stop career centers in the US state of Georgia. One-stop centers offer a full array of employment services including UI, core employment service (ES) like job interview referrals, intensive ES like counseling, and referral to job training. FDSS divides job seekers into finer groups than does WPRS. In each region of the state FDSS uses statistical models to group job seekers into five employability groups based on observable characteristics using administrative data. For each of the employability groups a different ranking of core and intensive ES are presented based on the reemployment success of recent service users. Similar rankings are given by employability quintile for alternative types of job training (remedial skills, job skills, and on-the-job). However, FDSS differs from WPRS in that it is simply a tool to help frontline staff make referral decisions to services, no sanctions result from choosing service patterns other than those ranked highest.

Uses of profiling in Canada

During the 1990s the Canadian federal government invested in developing the Service and Outcome and Measurement System (SOMS) (Colpitts 2002). SOMS was designed to determine the most effective employment service for a job seeker with particular characteristics. SOMS was an impressive technical system, but it was never accepted by frontline staff and was never implemented by the national system of Canada Employment Centers (CEC).

In 2002 the Canadian province of British Columbia (BC) adopted a subjective profiling technique to identify which new recipients of income assistance (IA)–cash welfare payments–were at greatest risk of long term IA receipt (British Columbia Ministry of Human Resources 2002). The methodology has been validated by independent psychometricians (R. A. Malatest and Associates 2002). The BC method involves two steps, both of which are subjective. The first is called the "Employability Screen" which has seven simple questions with categorical answers for each question. Each answer is assigned a point score and the scores are totaled. The second stage is called a Client Employability Profile, and it is computed for those with scores exceeding 14 on the first part. This second stage results in a grouping of clients into three

groups: those requiring little assistance, those requiring modest assistance, and those requiring intensive assistance. A customized service plan is then set depending on the category.

Elements of proposed profiling for Russia

The proposed methodology for early identification of customers of the public employment service in the Russian Federation who are most likely to suffer long term joblessness has two phases. The first phase is based on a statistical model estimating "employment potential." An objective assessment for an individual is produced by evaluating a statistical model reflecting the patterns of joblessness experienced by recent customers of the public employment service. Since data are not available on some characteristics of job seekers, such as prior occupation, rules are proposed to subjectively add factors directly to the statistically estimated model. Adding the result from the objective assessment with the subjective factors yields an "employment potential" score. A completely subjective approach to determining "employment potential" is offered in the second to last section of the paper. It is suggested that this second approach would totally obviate the need for statistical analysis of employment patterns of job seekers with differing characteristics and types of employment assistance.

The second phase is a subjective assessment of "employment motivation" based on two questionnaires, one applied to the job seeker and the other answered by frontline staff in the local employment bureau who has knowledge of the jobseeker's activity during the first ten days since registration. A point value is assigned to each possible response to questions. The points are added and combined from the two questionnaires to yield a score indicating "employment motivation."

Three separate categories of job seekers are set based on each of the two indices. To assign customers to employment services a three by three matrix is set up based on these groupings. Each job seeker is placed into one of nine cells based on their employment potential and employment motivation. The frontline staff can override the assignment to one of the cells based on personal professional assessment of employment potential and motivation. A suggested list of services is provided for each of the nine cells.

General Comments on the Proposed Russian Methodology

The proposed methodology is an ambitious process for assigning job seekers to employment services in local labor bureau office in the Russian Federation. It is a thought provoking plan, however little justification is given for the strategic choices made in setting up the system, and there is no evidence provided to validate the plan.

The profiling models used in North American employment programs have modest aims. They are instruments for guidance and rationing. Profiling applications in Canada and US all recognize that the tools are an imprecise way of distinguishing groups of customers, but in all cases there has been thorough efforts to validate the profiling mechanisms. Furthermore, an underlying principle of profiling in North America is that the level of services provided is directly related to the estimated probability of long term unemployment.

For the Russian proposal, the development and validation of the statistical model forecasting "employment potential" is not well documented. Among the alternatives presented in Table 1 there is no methodology offered for choosing a preferred model, or reasons given why other models were not tried. The coefficient of determination is a poor criteria for choosing a model to be used for profiling. What is important is how well the model actually predicts outcomes for clients. This is best validated by "in period" forecasts, wherein a random group from the estimation sample is reserved during estimation, and then later used to validate the forecast accuracy.

The discussion on methods for converting statistical model coefficients into point valuations is completely erroneous. The proper methods for forecasting an outcome and computing marginal effects from a logit model are presented in the page specific comments below. There is no reliable way to manually add additional factors *ex post* to the forecast result of a statistical model. The properties of such a system are intractable.

The Russian proposal also offers a completely manual and subjective approach to estimating "employment potential." It is asserted to be similar to the method used in British Columbia (BC), Canada. However, the Canadian method has been validated (as discussed below), and is applied to a much more limited service delivery choice: whether job search assistance should be provided to income assistance recipients.

The Russian proposal for assessing motivation for job seeking is entirely subjective and should be validated. It is impossible to assess the accuracy of this method without further investigation.

The Russian proposal for selecting active labor market programs based on categorizing job seekers into one of nine groups based on three levels of employment potential and three levels of employment motivation is arbitrary. It was not based on analysis of the effectiveness of each of the alternative services for job seekers in the different groups of employment potential and motivation. Subgroup analysis of the data on active labor programs used to estimate models summarized in Table 1, may provide some guidance.

Contrary to the Russian proposal, which recommends concentrating employment program efforts on "the intermediate and not extreme groups," the North American profiling systems for targeting services tend to concentrate services on those at greatest risk of long term joblessness.

Statistical profiling models should have modest uses. They are applied at the local level. A ranking of profiling scores at the local level is produced, and simple service referral decisions should be made based on the ranking. In the WPRS system in the US, profiling is used to ration limited spaces in intensive job search assistance, with those most likely to be long term unemployed given the slots for participation. In BC Canada the assessment is used to place customers into one of three groups: require little assistance, require modest assistance, require intensive assistance.

Application of profiling should not be applied either too strictly or divided too finely. Such uses are simply not justified given the sampling errors of the model forecasts. Furthermore, if profiling is based either wholly or in part on subjective factors the implicit sampling errors underlying referral rules will be compounded.

Page Specific Comments on the Russian Proposal

p. 1. Third line. Expand the statement to read: ...basis of individual characteristics of the unemployed person and local labor market. Note: local labor market conditions can be captured simply by indicator variables. (However, their methodology envisions estimating separate models for each local labor market.)

P. 2. Section 2. Let the section begin: This section discusses the estimation of a statistical model to assess....

P. 2. Several models of "employment potential" are put forth. Be more clear about what model is recommended. Models with two alternative dependent variables are considered. Be clear about what the outcome of the statistical model with the dichotomous dependent variable is, e.g., binary (1, 0) setting the threshold at 4 or 6 months. Note that when the dependent variable is continuous, the logarithm of the duration of unemployment is chosen, but in many cases this is a censored variable. That is, the true duration is not observed, only the time to the current date. In this case an estimator that accounts for censoring, such as tobit, may be appropriate.

P. 2. Add a footnote to explain exactly what "statistical allowances are more accurately complied with" by using the logarithm of duration unemployed as the dependent variable in a model estimated by ordinary least squares instead of the level of duration.

P. 2. Bottom paragraph commenting on survival models should be removed, it is irrelevant to the presentation.

P. 3. Delete first sentence, or move the statement about using Stata to a footnote.

P. 3. Revise the second sentence to read: The model predicting long term employment based on individual characteristics is the basis for a computer algorithm that returns a probability estimate for the individual to the frontline staff in employment bureau offices.
p. 3. It is a good suggestion to involve regional academic centers in a plan for future updates of profiling models.

P. 3. The last sentence of the first paragraph is unclear. It does not add anything to the explanation. The sentence may be removed.

P. 3. Section 2.2. It is a very good suggestion to encourage standard software nation-wide for administration of employment services. This is a serious obstacle to using an automated profiling methodology. It also hinders nation-wide performance monitoring and other evaluation studied.

P. 3. Section 2.2. Occupation is an important factor in profiling models. It should be possible to introduce a simple occupational coding system based on high level occupational groupings. One possibility is to apply the International Standard Classification of Occupations (ISCO) established by the International Labor Office in 1988.

ISCO-88 consists of four levels of categorization. These are defined as:

10 major groups
28 sub-major groups (subdivisions of major groups)
113 minor groups (subdivisions of sub-major groups)
377 unit groups (subdivisions of minor groups)

The 10 major groups would be sufficient to add information to the profiling model. The 28 sub-major groups would be even more informative if possible.

P. 3. In section 2.3, it is unclear where there is a prior reference to "choice of statistical and econometric models." Why was Voronezh chosen. Is it representative as a typical labor market? ... typical for computer methods of administrative data on employment services? ...was there a good personal connection by the project developers? ...other reasons? Please explain in the text.

P. 3. Section 2.3. "The data refer to unemployed who were recorded in 2000 as having been dropped from the register and those registered at the end of the year." A more appropriate sample selection strategy would be to choose the entire sample inflow to the register during some period of time and to check the status of these people after 4 or 6 months (depending on the definition of the dependent variable). Always select an inflow sample for analysis. Selecting an outflow sample introduces a sample selection bias.

p. 3. What inducements are there to register as a job seeker or as unemployed? Are there entitlements to unemployment compensation or access to public health services? These factors may explain entry and duration of registration. People may return to informal sector employment but wish to maintain access to health services.
p. 3. Controlling for involvement with active labor market programs (ALMPs) is a delicate matter. At what time in a job seeker's unemployment spell will the profiling model be evaluated? If it is before usual entry to ALMPs, they will not be relevant when assessing the model. However, it may be important to control for ALMPs in estimation to avoid omitted variable bias in estimating other parameters.

p. 3. When you say "sector" to which last place of employment belonged, do you mean industry group? For example, agriculture, construction, manufacturing, retail, services,...? If yes, then industry group would be a better description. It would also be useful to list the 20 industry groups used, perhaps in a footnote.

p. 4. It should be noted that profiling models in the United States are prohibited from using gender, age, or race as factors because it would result in systematic discrimination that is legally prohibited (Wandner 1994).

P. 4. There is a statement that certain data (training cost, training duration, status on second review) were not included in the model. Profiling models should be minimalist. That is, variables should be added to the specification only if they add significant explanatory power. Since the model will be evaluated early during the registered unemployment spell, such factors are probably irrelevant to predicting long-term unemployment. Though they may be relevant to proper estimation of other relevant parameters, thereby avoiding omitted variables bias.

P. 4. Seasonal factors may improve the predictive accuracy of a model estimated on a sample drawn over a 12 month period, however they will add nothing to the ability of a model to rank job seekers entering local employment bureaus during a given week or month (Black et al. 2002).
P. 4. Replace the first sentence in the fifth paragraph. The models have one function, to forecast the likelihood of long term unemployment.

p. 4. Discussion of model factors:

Age – is prohibited from use in the Worker Profiling and Reemployment Services (WPRS) model in the US. In the British Columbia (BC) Canada methodology and the Frontline Decision Support System (FDSS) system in the US it is used only to identify youth.

Education - is a standard variable in profiling models. Usually, higher educated persons have a higher probability of long term unemployment. Most high educated persons change jobs without ever visiting the public employment service (PES). High educated persons who seek assistance from the PES typically have exhausted prospects from informal networks, and are at great risk of long term unemployment.

Job Tenure - length of service, patterns are usually similar to education. Higher tenure means a greater chance of long term unemployment.
Work Experience - the effect of more lifetime work experience on the chance of long term unemployment is uncertain.

Reason for Separation - effects of dismissal, redundancy or other reasons for separation are uncertain.

Number of Dependents, Marital status, Firm ownership type, Separation for cause - these variables could operate in either direction.

ALMP participation - Inclusion of these variables in estimation of the model may be necessary to correctly estimate other model parameters, however if profiling is done early after registration with the public employment service, most profiled clients will not have used these services and they will not be directly relevant to profiling. As discussed below, however, analysis if this data may provide a guide for service referral decisions.

Prior industry - A set of prior industry dummy variables is an important component of profiling models.

Geographic region of employment - these variables may affect how models fit, but they are not relevant to the ranking of job seekers in any particular local office.

Gender - is prohibited from use in the Worker Profiling and Reemployment Services (WPRS) model in the US. It is hard to imagine public acceptance of a profiling model that treats men and women differently.

Using r squared as an indicator of model performance - is not an adequate indicator. As discussed below, a portion of the analysis sample should be randomly selected and put aside before estimation and then later the reserved sample can be used for validating the forecast accuracy of the model. Simply compare the model forecast outcomes on the reserved sample to the actual outcomes.

p. 5. What is meant by the "limit effect?" The subsequent description in the sentence is that of a marginal effect, which is distinct from the parameter estimate of a logit model.

P. 5. R squared is an indicator of the proportion variation in the dependent variable explained by the right hand side variables in the model, but it is not a good indicator of forecast accuracy. That should be assessed by out of sample validation. That is, from the sample for estimation, a subset of observations should be randomly selected and reserved for assessing forecast accuracy. For example, randomly select 10 percent of the sample, estimate alternative specifications using the 90 percent and evaluate the forecast accuracy of the model using the 10 percent sample reserved.

p. 5. It is not clear what the dependent variable is in the simplified model, there is poor justification for the simplified model. With electronic computer computations, why must the model be simpler? Which model is the system based on, probability of long term unemployment or duration of joblessness? Also note that the duration of joblessness is measured inaccurately because of censoring. At any point in time we observe only truncated durations of joblessness for those still looking for work. This censoring must be accounted for in estimation.

p. 7. Section 2.4 is irrelevant. This is not what is being done in British Columbia, Canada. References on that methodology are R. A. Malatest and Associates (2002) and British Columbia Ministry of Human Resources (2002). The British Columbia (BC) method is relevant to the discussion of completely subjective scores as proposed in section 4 of the report, but there is nothing in the BC method that suggests a way to combine scores from a statistical model together with subjectively assigned points. Furthermore, there is no need to convert the meaning of individual parameter estimates into anything other than marginal effects. The most important use of statistical models for profiling is to evaluate the model to compute an outcome valuation for an individual new customer so that customers can be ranked on these scores and referred to services.

For models estimated by ordinary least squares, marginal effects and prediction of the outcome for an individual are direct. The methodology for computing marginal effects of parameters in a logit model and for estimating the probability of an outcome for an individual using a logit

model follows:

When a model such as:

$y = a + b1x1 + b2x2 + b2x3 + ... + bnxn + e$

is estimated by logit, where the dependent variable is binary, i.e. y = 0 or 1.

Predicted probability of y = 1 for a given person, observation i, Prob (yi)

Let: sum XiB = A + B1x1i + B2x2i + B3x3i + ... + Bnxni

where A, B1, B2, B3, Bn are the logit parameter estimates
and B is the matrix of all logit parameter estimates
x1i, x2i, x3i, ... , xni are the values of the variables for person i
and Xi is a matrix of all variables (including a constant) for person i

Then the predicted value of y = 1 for person i is computed by the formula:

Prob (yi) = (1/(1+exp(-XiB))

where exp is exponentiation with the base being the inverse of the natural logarithm or e.
Logit parameter estimates are not interpreted as marginal effects of the variable on the outcome. A transformation is required to produce marginal effects. The transformation differs depending on whether the independent variable, x, is continuous or binary (1, 0). If x1 is a continuous variable like years of work experience, then the marginal effect of x1 on y is computed as:

(mean x1)*(1 - (mean x1))*B1, where B1 is the logit parameter estimate on x1.

If x1 is a discrete variable, then the marginal effect of x1 on y is computed as:

[r*exp(B1)/(1 + r*exp(B1))] - x1, where r = (mean x1)*(1 - (mean x1))

Technical background for these computations is explained in Maddala (1983, p. 23). An example in the context of profiling for long-term unemployment is given in Eberts and O'Leary (1996) which is available at http://www.upjohn.org/publications/wp/9641.html

p. 8. Section 2.5. To refine the profiling process by adding factors beyond those with data available is a difficult task. Weights cannot be simply concocted and subjectively added to the forecast result of a statistical model. There is no simple way to calibrate the scale of added factors. On page 12, there is an explicit reluctance to combine the results from two subjective indices "because we did not know how to weight these indicators." For this essential reason, results from the statistical model should not be combined with subjective add factors.

p. 8. Ideally categorical information on prior occupation should be added to the data for estimation. The WPRS profiling models in the US typically use ten major occupational groups.

As discussed above, adding ISCO-88 categorization based on either 10 major groups, or 28 sub-major groups (subdivisions of major groups) could be very valuable to the modeling exercise.

p. 8. No scientific basis is offered for the assignment of add factors concerning "how easy it is to find work for an unemployed person with a certain profession." Such assessments require evidence on occupational labor demand. Translating the importance of prior occupation into reemployment probability requires statistical estimation based on historical data. Subjective assessments of labor demand by occupation cannot be meaningfully added to an index of the probability of long term unemployment based on a statistical model. The proposed methodology offers no justification for adding a point or subtracting half a point. The rule is simply imposed *deus ex machina*.

p. 8. A separate employment potential index could be constructed, but it should not be a simple add factor addendum to a statistical model. The proposed factors are extremely subjective and have not been validated by any objective assessment. If referral to services is based on these subjective assessments it is an opportunity for front line staff to "game the system." That is, to make referrals as they choose, regardless of the true likelihood of long term unemployment.

p. 8. The method of assessing motivation for job seeking based on two questionnaires. The first list of questions are asked of job seekers to assess motivation for job seeking. The second list of questions are answered by front line staff based on the observed job search activity during the first 10 days since registration as unemployed.

p. 9 - 12. A subjective index of motivation for employment is possible, but it should be properly validated and modestly used. This point is asserted on page 12 "testing in regions will help to define more accurately the critical values for the motivation index," however, no validation exercises have been attempted.

p. 12-13. Section 5. I would be interested to see the evidence from Chelyabinsk and Karelia about what services are appropriate for customers with different characteristics. There is no statistical model used in Canada for assignment to services. The national government in Canada developed and tested a Service and Outcome and Measurement System (SOMS) that was a statistical model based referral to services, but it was never implemented (Colpitts 2002). The system used in the province of British Columbia, Canada is entirely subjective (British Columbia Ministry of Human Resources 2002), however it has been scrutinized and validated by independent psychometricians (R. A. Malatest and Associates 2002). Statistical models are the basis of the WPRS systems for early referral to job search assistance in the US (Dickinson et al. 2002). A statistical model is the basis of service referral in the FDSS system being pilot tested in the US state of Georgia (Eberts and O'Leary 2002).

p. 12. You assert that for "low potential, low motivation" customers "it makes sense to just pay them unemployment benefits" and instead focus active services on other clients. This is exactly the opposite from the rule applied in the WPRS system in the US where those identified at risk of long term unemployment are quickly referred to job search assistance to promote return to work and conserve UI payments.

p. 13. You propose “the test of the claim to unemployment (benefits is)... recommended for those who have no motivation to seek employment” Several evaluations in the US have repeatedly found that an actively enforced and verified work test applied to unemployment compensation recipients promotes return to work and saves unemployment compensation benefits (O’Leary 2002).

p. 13. Several services appear in more than one cell. Does the ordering of services within a cell suggest the ideal order of participation for customers classified in that cell?

P. 13. Section 6. Beginning with the second line, let it read: methods in the form of a computer model in the Microsoft Excel spreadsheet. Frontline staff in local employment bureaus would simply enter the appropriate data for an individual customer using the graphical user interface. On the basis of data entered, the potential motivation and list of recommended ALMPs are automatically determined. Among other factors, the computer model includes:

p. 14. Top line. Let it read: At the testing stage an additional field is used. It is designed to reflect why the employment service officer disagrees with the proposed ratings;
The recommended choice of programs for a given unemployed person, at the testing stage, can be from the list of programs, or from an altered list. All the changes....

P. 14. Because the methodology used is wrong, the third sentence of section 7 should be removed. That is, remove the following: “The points assessments used in the questionnaire are converted estimates of the coefficients of the model discussed above.”

P. 14. In the second paragraph of section 7, replace the word “filled” in several places with either “answered” or “completed.”

P. 14. In the fourth paragraph of section 7, let the second line begin: ...sum up the number of points representing answers to all questions. The number of points awarded for each response is indicated in the far-right gray-colored column.

p. 14. In the fifth paragraph of section 7, there is no analytical or statistical basis for the assertions regarding the proportion of customers that will fall within any range of scores from the proposed subjective rating schemes.

p. 16. In the second paragraph of section 8, there is an assertion that “the procedure is scientifically valid.” There is insufficient documentation to support this claim.

p. 16. In the fourth paragraph of section 8, is an important point “problems stem from the lack of a standard format for entering and coding data from personal registration cards.” As mentioned above, differing formats for registration and saving personal data hinder performance monitoring and future net impact evaluation studies. Information should cover important items, be recorded in a consistent manner, and be required for all service users.

p. 16. In the fifth paragraph of section 8, the valid points are made that models should be revised periodically and that regional capacity at universities or research institutes should be used to do

this. These are important points, however, there need not be separate models for each local area, and labor markets probably don't change sufficiently in one year to justify revision. Since it's the ranking of clients at the local level that matters for referral to services, regional (oblast level) models should be sufficient. Regional models revised every 3 or 4 years would probably be adequate.

p. 16. It is an important point to make that objective profiling based on statistical models should be used judiciously by professional employment services staff when making referral decisions. Some automated referrals are practical, such as to job search assistance that has been shown to be generally effective and low cost, but profiling rankings are better used as a tool rather than a rule when making choices among a complex set of services.

References

Black, Dan, Jeffrey Smith, Miana Plesca and Suzanne Plourde. 2002. "Estimating the Duration of Unemployment Insurance Benefit Recipiency." Final Technical Report. Contract Number UI-10908-00-60. Washington, DC: U.S. Department of Labor (February). http://www.glue.umd.edu/~jsmithz/papers/FinalETA110102.pdf

British Columbia Ministry of Human Resources. 2002. *Employment Assistance Manual*. Victoria, BC, Canada: Author (September). www.mhr.gov.bc.ca/PUBLICAT/VOL1/index.htm

Colpitts, Terry. 2002. "Targeting Reemployment Services in Canada," in *Targeting Employment Services*, Randall W. Eberts, Christopher J. O'Leary and Stephen A. Wandner, eds. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. http://www.upjohn.org/publications/titles/tes.html

Dickinson, Katherine P., Paul T. Decker, Suzanne D. Kreutzer, and Richard W. West. 1999. "Evaluation of Worker Profiling and Reemployment Services: Final Report." Research and Evaluation Report Series 99-D. Washington, DC: Office of Policy and Research, Employment and Training Administration, U.S. Department of Labor. http://wdr.doleta.gov/opr/fulltext/document.asp?docn=5885

Dickinson, Katherine P., Paul T. Decker and Suzanne D. Kreutzer. 2002. "Evaluation of WPRS Systems" in *Targeting Employment Services*, Randall W. Eberts, Christopher J. O'Leary, and Stephen A. Wandner, eds. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. http://www.upjohn.org/publications/titles/tes.html

Eberts, Randall W., and Christopher J. O'Leary. 1996. "Design of the Worker Profiling and Reemployment Services System and Evaluation in Michigan," Staff Working Paper 96–41. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research (April). http://www.upjohn.org/publications/wp/9641wp.html

Eberts, Randall W., and Christopher J. O'Leary. 2002. "A Frontline Decision Support System for Georgia Career Centers," Staff Working Paper 02-84. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research (July). http://www.upjohn.org/publications/wp/0284wp.html

Maddala, G. S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. London: Cambridge University Press.

Messenger, Jon, Carolyn Peterson-Vaccaro, and Wayne Vroman. 2002. "Profiling in Self-Employment Assistance Programs," in *Targeting Employment Services*, Randall W. Eberts, Christopher J. O'Leary, and Stephen A. Wandner, eds. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. http://www.upjohn.org/publications/titles/tes.html

O'Leary, Christopher, Paul Decker, and Stephen Wandner. 1998. "Reemployment Bonuses and Profiling," Staff Working Paper 98-51. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research (January). http://www.upjohn.org/publications/wp/9851wp.html

O'Leary, Christopher J. 2002. "Effectiveness of Labor Exchange Services," presented at the conference "Job Training and Labor Exchange in the United States" jointly sponsored by the W.E. Upjohn Institute and the U.S. Department of Labor, September 19-21, Augusta, Michigan.

R. A. Malatest and Associates. 2002. "Summary Report: Review of the Service Screening Tool" prepared for the British Columbia Ministry of Human Resources. Victoria, BC, Canada: Author (June).

Wandner, Stephen A., ed. 1994 "The Worker Profiling and Reemployment Services System: Legislation, Implementation Process and Research Findings." Unemployment Insurance Occasional Paper No. 94-4. Washington, DC: UI Service, Employment and Training Administration, U.S. Department of Labor.

Wandner, Stephen A., and Jon C. Messenger, eds. 1999. "Worker Profiling and Reemployment Services Policy Workgroup: Final Report and Recommendations," Office of Workforce Security Paper. Washington, DC: U.S. Department of Labor (February). http://wdr.doleta.gov/opr/fulltext/document.asp?docn=6074*Project "Development of Unemployed Profiling Procedures as a Basis for Planning and Implementation of Active Labor Market Programs*

Procedure for Profiling Unemployed Citizens in Russia

1.Concept of Methodology of Profiling Unemployed Citizens in Russia

By *profiling* we mean a system of assigning unemployed citizens to categories (groups) in accordance with the risk of prolonged unemployment. In this case the risk was assessed on the basis of the individual characteristics of the unemployed person, reflecting the potential and motivation for employment, with a view to identifying a set of active labor market programs (ALMP) that best enhance the chances of employment for those groups that merit assistance on the part of the employment service, in terms of economic efficiency and considering the limitations of resources.

In other words, profiling is a procedure for identifying problem categories with a high risk of prolonged unemployment *at an early stage* and involving them in special programs from the first weeks that they have been entered in the register of unemployed. Special programs are designed to focus the efforts of the unemployed and employment service personnel on making the unemployed person more competitive on the local labor market and have been tailored to the individual features of the unemployed person.

Based on international experience, and considering the Russian practices of work with unemployed people (including the nature of the information base), we propose, in the first phase of profiling (when assessing the risk of prolonged unemployment) to use a statistical model adjusted on the basis of the opinions of experts in the employment service, and in the second phase (when selecting programs for every type) to rely on the opinions of employment service experts. We suggest that the profiling procedure be carried out at the phase when citizens are registered as unemployed and not when they first come to the employment service (the time of primary registration)[1].

The proposed procedure for profiling comprises the following elements:

- Methods of estimating the statistical model of employment potential based on individual data from the register of unemployed in the region, with an example of building and estimating a model based on the data from a pilot region;
- Methods of converting statistical model coefficients into point-scale assessments that can be conveniently used by employment agencies in assessing the employment potential of a concrete unemployed person;
- A list of additional questions to an employment center worker designed to fill the gaps in the statistical model caused by lack of immediately available information on the parameters determining the chances of employment suitable for analytical processing;
- Two questionnaires to assess the motivation of unemployed persons to find a job. The first questionnaire is a list of questions asked by the profiling specialist of the unemployed person in order to assess the motivation to seek employment. The second questionnaire is a complement to the first one: it is a list of questions answered by the profiling specialist, based on the information at his disposal on the work of the unemployed with vacancies in

[1] A large share of initially registered citizens find a job within the first ten days and never get to the stage of becoming registered as unemployed.

the ten days following first registration with the employment service;

- Methods of deriving summary assessment of motivation towards employment based on the two above-mentioned questionnaires;
- Methods of identifying groups of unemployed on the basis of the derived assessment of the risk of prolonged unemployment which have two dimensions: assessment of employment potential and assessment of motivation to seek employment[2].
- Schemes of selecting an ALMP menu for selected groups of unemployed. Prescription of ALMPs for subgroups is based on the experience of Russian employment services.

The Consultant has also prepared a mock-up of an automated profiling method which is a macros in the EXCEL software program, formalizing the profiling procedure. This permits, on the one hand, presenting the procedure of profiling in a form that can be conveniently used by employment officers and on the other hand, does not require additional software support[3]. In addition to the automated procedure we have developed a questionnaire for determining the employment potential that lends itself to manual processing.

2.Method of Estimating Statistical Model of Employment Potential of Unemployed Citizens

This paragraph discusses the building of a statistical model to assess the probability of a prolonged period of unemployment and/ or the duration of the period on the basis of individual data for registered unemployed (the basis of the first phase of profiling).

2.1.Choice of Statistical and Econometric Models

The dependent variable in the model used to forecast the probability of prolonged unemployment is defined as being out of work for a period exceeding 4 months[4]. A binary variable equals 1 if a person is registered as unemployed for more than 4 months, and 0 if less than 4 months. The analysis of such discrete dependent variables calls for logistical regression and other methods of analyzing limited dependent variables.

To analyze the duration of unemployment a linear regression model is used in which the dependent variable is the logarithm of the duration of unemployment. The advantage of this model compared with one which analyzes (non-logarithmic) time out of work consists in that, on the one hand, the weight of remote observations (unemployed registered for several years) is diminished and, on the other hand, the resolution capacity of the model for short periods (from several days to several weeks) is increased. In addition, certain statistical allowances used to derive formulas of linear regression assessment are more accurately complied with.

There is yet another class of econometric models called models of survival, withdrawal and

[2] In the future regular gathering of information on the assessment of an unemployed person's motivation by an employment center worker may make it possible to include motivation assessment in the body of variables in the statistical model and, consequently, to bring two dimensions of prolonged unemployment risk – potential and motivation -- into a single index.

[3] In the future the procedure of profiling is to be integrated into the professional software of the employment officer ("Employment and Control") developed by the Labor Ministry.

[4] The experience of Russian employment services shows that the critical period in looking for a job is about four months. As a rule, this time is sufficient for highly motivated workers for whom there is demand in the labor market to find employment; those who stay out of work for more than four months may need additional assistance in the form of ALMPs. However, that critical value can easily be altered, depending on the features of the region, to six months, for example.

refusal, intended specifically for analyzing such data as the time of transition from one state to another (in this case from being out of work to being in work, or refusal to look for work). The object of study for such models is the instant probability (called risk) of transition (finding work) at a given moment of time on condition that the subject is in a different state (without work) during the said period. Such models make it possible to take account of individual characteristics of the unemployed, as a rule in the shape of co-multipliers that increase or diminish the "risk" of change of status. Such models are not used in this phase of the project although they are of undoubted scholarly interest.

The analysis uses the Stata package, a statistical and econometric package that best conforms to the tasks of analysis of individual data by the above methods. The output of the analysis procedure will be a system of weight coefficients for individual characteristics which are then used in the software in the employment official's workplace. Such coefficients should preferably be subject to regular updating in accordance with changes in the economic situation in the regional labor markets. For this work to be competently accomplished regional academic centers should be brought in. If difficulties arise at the first stage it is possible (though not desirable) to stay at the level of an economic region or a federal district.

2.2.Presentation of Data

Employment services use several standard interfaces of data bases including the systems Catharsis (Karelia) and Assistance (Omsk, Chelyabinsk) as well as some specialized products developed in the regions (Voronezh). All these data are presented in DBF format, the most popular data base format.

It should be noted that lack of a unified format for data retrieval and encoding may be a further hindrance in introducing the proposed profiling methods in some regions at the initial phase. We also recommend that employment services pay more attention to unifying the data of various administrative entities within large cities and whole regions.

We suggest one of two approaches to deal with the problem for lack of uniformity of data:

- to unify the data bases by developing common requirements to the data base, a list of variables and formats of presentation, in particular, geared to the tasks of profiling;
- to set before employment service IT workers the common task of developing software for searching and using the necessary information on the basis of existing data bases and systems of controlling them.

Another substantial drawback of existing data bases for unemployed persons is the shape in which information is gathered on types of occupations/ professions which are encoded in a format that is unfit for analytical processing.

2.3.Example of Building a Model of Unemployment Duration

This section describes a case which assess the models referred to above in the section "Choice of Statistical and Econometric Models", with reference to data for the city of Voronezh and Voronezh Region.

The data used refer to unemployed who were recorded in 2000 as having been dropped from the register (because they had found employment, retired, moved out of the area, etc.) and those registered at the end of the year (who had not yet found a job). The following individual characteristics were used:

- gender;
- age (linear and quadratic terms);

- family status;
- number of dependents;
- educational level;
- length of employment: total and in the last place of work;
- cause for dismissal;
- involvement in ALMPs (indicator variables which have the value of 1 if the unemployed person took part in one of the ALMPs: public works, Youth Practice, New Start, occupational counseling, Club of Job Seekers, retraining);
- area of the region in which the unemployed person resides[5]
- sector to which the last place of employment belonged (20 titles).

For some variables cross terms with gender were designed indicating the difference between model coefficients for men and women[6].

Some characteristics (such as the level of skills, wages in the previous place of work and characteristics of retraining programs – cost, duration, status on second review) have not been taken into account in this model because such data were inaccessible or incomplete. A possible partial answer to the problem of inaccessibility of essential data on employment potential is discussed in section 2.5.

In some regions unemployment has a pronounced seasonal character. To take account of these features dummy variables corresponding to quarters of the year may be, and in fact have been, introduced for some pilot regions.

The technical results of regressions based on data for the Voronezh Oblast are contained in Table 1. Below is our interpretation of the results.

The results of regressions have two functions: analytical and prognostic. In terms of analysis of the characteristics related to the duration of being out of work, one must note the statistical significance[7] of the following variables are of note:

- age of person – the dependence is linear;
- the negative coefficient for the "year of birth" variable (see definition in note to Table 1) confirms that people of older ages have more difficulty finding a job;
- education – specialized education and special skills which the worker seeks to use to the maximum degree tends to increase the time of job seeking;
- length of service in the last place of work: the longer the period the more difficult it is to find a new job (because the worker has settled in the habits peculiar to the last place of work which he/ she finds it hard to apply working for another employer);
- overall length of service – despite with the negative effect of age, more experienced workers find it easier to find work. Continuous employment may have an effect: those who had a break in service and dropped out of the labor market for a long period have more difficulty finding a job;
- cause of dismissal – workers laid off as part of staff cuts (made redundant) or liquidation of

[5] The problems of low mobility and regional segregation of labor markets in Russia are well known.

[6] Gender differences in the behavior in the labor market may be taken into account by assessing the model *separately* for males and females, which has been done for the republic of Karelia, the Chelyabinsk and Omsk oblasts

[7] Statistical significance describes the probability of observing a given relationship by chance, even though when the given factor has in fact no real effect. Low probability is associated with low likelihood that the relationship is observed by chance. The commonly accepted levels of significance of 1% and 5% are marked with asterisks in the table.

an enterprise tend to stay on the register of unemployed longer. One explanation is that the employer pays them unemployment benefit for a longer period (the first three months if registered with the employment service in the first two weeks after being laid off) which is a disincentive to real search for work;

- number of dependents – a large family as a rule forces a worker to look for a higher paid job which is more difficult to find;
- involvement in ALMPs is connected with more prolonged search for work, probably because unemployed people who have more difficulty in finding work are sent to ALMPs[8];
- retraining greatly enhances the chances of finding work;
- such variables as marital status and the form of ownership of the enterprise in which the person was previously employed are statistically insignificant;
- interestingly, workers dismissed for disciplinary offences take less time to find a job (the variable is insignificant in models for probability of looking for work for more than 4 months). Such individuals may be more active in the labor market;
- There is a pronounced sectoral and geographic effect: workers in different sectors and from different regions find work with different speed[9];
- Women have more problems finding work. This is seen not only in the overall gender effect (described by the variable "gender" in the third line of the table) but in the fact that women have more difficulty finding a job in each of the sectors enumerated. But the age effect is less pronounced, as is the effect of the number of dependents.

The numbers in columns (1) and (2) of Table 2 describe changes (limit effect) in the probability of prolonged unemployment if the value of each of the above variables increases by unity. In the remaining columns the values of coefficients are given. For log-linear models in columns (3) and (4) the values of coefficients show the average percentage by which the time of job seeking by worker with a certain characteristic differs compared with the "zero" level. The multiple determination coefficient R^2 in the bottom line shows that the model explains about 16% of the total spread of the job seeking time logarithm. It means that the prognostic power of the model is comparatively small. Unrecorded individual features of workers provide a spread that is about double the variability predicted by the model. This shows that the statistical model alone cannot be regarded as sufficient in profiling. The final column contains coefficients of a simplified probit-model for the probability of prolonged unemployment (over 4 months).

The prognostic function of regressions consists in building index functions to characterize the probability of prolonged stay out of work and/ or job seeking (depending on the model used). To avoid cumbersome computations, the simplified model (column 5) drops statistically insignificant variables as well as multi-category factors such as "region" and "sector"[10].

[8] It should be noted that the assessment of ALMPs indicator coefficients may be distorted because of the way people are selected for these programs. The empirical level of value may be inaccurate and the interpretation of coefficients mistaken.

[9] Information is lacking for these two factors in about one half of the cases, so the assessments of coefficients (shifts of categories) for these variables are probably not accurate enough.

[10] This allows inaccurate specification of the model because previous analysis has shown the significance of both factors.

Table 1. *Assessments of parameters of unemployment duration models.*

	(1) Prolonged unemployment	(2) Prolonged unemployment	(3) Ln(time out of work)	(4) Ln(time out of work)	(5) Simplified model
Age	0.0092	0.0083	0.0199	0.0175	0.0231
	(0.0004)**	(0.0006)**	(0.0007)**	(0.0009)**	(0.0012)**
Age^2	-0.0043	-0.0007	0.0061	0.0173	
	(0.0033)	(0.0048)	(0.0054)	(0.0080)*	
Gender: female	0.0987	0.1653	0.2110	0.4030	0.3306
	(0.0067)**	(0.0302)**	(0.0111)**	(0.0500)**	(0.0221)**
Family	0.0061	0.0059	0.0163	0.0153	
	(0.0071)	(0.0071)	(0.0117)	(0.0117)	
Vocational education	0.0350	0.0351	0.0714	0.0719	0.1029
	(0.0069)**	(0.0069)**	(0.0114)**	(0.0114)**	(0.0172)**
Higher education	0.0556	0.0561	0.1389	0.1389	0.2330
	(0.0080)**	(0.0080)**	(0.0134)**	(0.0134)**	(0.0314)**
Length of service in last place of work	0.0022	0.0021	0.0036	0.0034	0.0085
	(0.0008)**	(0.0008)**	(0.0013)*	(0.0013)**	(0.0020)**
Total length of service	-0.0072	-0.0073	-0.0145	-0.0146	-0.0153
	(0.0004)**	(0.0004)**	(0.0006)**	(0.0006)**	(0.0009)**
Dismissal for disciplinary offenses	-0.0704	-0.0661	-0.2241	-0.2083	
	(0.0427)	(0.0428)	(0.0712)**	(0.0713)**	
Redundancy	0.1652	0.1647	0.3909	0.3876	0.4743
	(0.0078)**	(0.0078)**	(0.0135)**	(0.0136)**	(0.0199)**
Number of dependents	0.0272	0.0429	0.0533	0.0971	0.1444
	(0.0049)**	(0.0083)**	(0.0081)**	(0.0139)**	(0.0185)**
Public works	0.2801	0.2810	0.5292	0.5309	
	(0.0100)**	(0.0100)**	(0.0215)**	(0.0216)**	
Retraining	-0.2606	-0.2609	-0.4070	-0.4031	
	(0.0485)**	(0.0486)**	(0.0818)**	(0.0818)**	
«New start»	0.3021	0.3020	0.5352	0.5349	
	(0.0174)**	(0.0174)**	(0.0372)**	(0.0372)**	
Club of Job Seekers	0.3057	0.3056	0.4268	0.4251	
	(0.0189)**	(0.0189)**	(0.0435)**	(0.0434)**	
Youth Practice	-0.0289	-0.0349	0.2450	0.2297	
	(0.1517)	(0.1523)	(0.2450)	(0.2448)	
Occupational guidance	0.0782	0.0789	0.1606	0.1634	
	(0.0137)**	(0.0137)**	(0.0225)**	(0.0224)**	
Government sector	0.1610	0.1565	0.0808	0.0701	
	(0.1301)	(0.1301)	(0.2123)	(0.2121)	
Private sector	0.1163	0.1118	-0.0157	-0.0256	
	(0.1322)	(0.1322)	(0.2124)	(0.2122)	
Female* number of dependents		-0.0206		-0.0574	-0.0666
		(0.0100)*		(0.0166)**	(0.0220)**
Female* age		-0.0017		-0.0044	-0.0031
		(0.0008)*		(0.0013)**	(0.0014)*

Female* age^2		-0.0018		-0.0090	
		(0.0068)		(0.0111)	
Female * vocational education		-0.0055		-0.0116	
		(0.0148)		(0.0243)	
Female * higher education		-0.0661		-0.1505	-0.1709
		(0.0172)**		(0.0281)**	(0.0371)**
Regions:	chi2(21) =37.63	chi2(21) = 37.95	F(23, 30602) = 2.47	F(23, 30579) = 2.46	
Empirical significance	0.0142	0.0131	0.0001	0.0001	
Sectors	chi2(21)=152.81	chi2(21) = 45.75	F(21, 30602) = 8.60	F(21, 30579) = 3.37	
Empirical significance	0.0000	0.0014	0.0000	0.0000	
Female		chi2(20) = 45.69		F(20, 30579) = 2.44	
Sectors		0.0009		0.0003	
Free term			4.5903	4.4462	-0.1166
			(0.0264)**	(0.0423)**	(0.0196)**
Number of observations	30676	30676	30666	30666	31271
R^2			0.16	0.16	

Note: Standard errors of assessments indicating approximate spread of values around the point assessment are in brackets. * : significant at 5% level; **: significant at 1% level. Columns (1), (2): maximum effects (average change of probability of long-term unemployment if the factor is valid). Columns (3)-(5): model coefficients.

2.4.Methods of Converting Statistical Model Coefficients into Point Valuations and Computing Potential for Finding Employment

The model of forecasting the probability of prolonged stay out of work may be presented in the shape of a points table (Table 2) which employment service workers can conveniently fill in, similarly to how it is done in Canada. Such a conversion is particularly convenient if the procedure of profiling is not automated, and the questionnaire is filled by the unemployed person to be subsequently processed by the employment specialist.

For recalculation into points, linear transformation is applied to statistical assessments of the model coefficients obtained[11]. This yields more "convenient" values of coefficients and convenient boundaries of predicted probability of prolonged unemployment: the boundaries of 25%, 50% and 75% correspond to the values of the new index equal to 0, 2 and 4.

Let us consider examples of converting coefficients into points. Point assessment of the ages of males is obtained by additionally multiplying the statistical assessment of coefficient at age 0.0231 (see column (5) of Table 1) by 3:0.0231*3=0.07. In calculating age points for females the value of the cross term (female * age) needs to be taken into account. That is, the summary assessment equals: 0.0231*3+(-0.0031)*3=0.06. All the other coefficients are similarly converted into points.

[11] The coefficients are additionally multiplied by 3 and their sum shifted by 2. The choice of multiplier and shift coefficient is based on the fact that the quartile (25%) of a standard normal distribution is approximately equal to 2/3.

Table 2. *Points assessment of individual characteristics.*

Characteristic	Registration data of unemployed	Coefficient for males	Coefficient for females	Result: product (2)*(3) or (2)*(4)
Age		0.07	0.06	
Gender		-0.80[12]	1.00	
Education:				
secondary		0		
secondary specialized		0.30		
Higher		0.70	0.20	
Length of service:				
Total		-0.05		
in the last place of work		0.03		
Made redundant/liquidation		1.42		
Number of dependents		0.43	0.24	

On the basis of the points table the points index of the probability of staying out of work for more than 4 months or the *employment potential* can be calculated. To do so it is necessary to introduce the personal data of the unemployed in the second column of Table 2 and multiply them by corresponding coefficients and add them up. For example:

- a man 30 years of age, no children, secondary education, total length of service 10 years, length of service in the last place of work – 1 year: 30•0.07-0.8•1+0-0.05•10+0.03•1+0.43•0=0.83. The potential of employment equals 0.83 which corresponds to a 25-50% probability of staying out of work for more than 4 months.
- A woman of 50, higher education, worked for 24 years in one and the same place, one child, dismissed as part of staff cuts:

50•0.06+1•1+0.20•1-0.05•24+0.03•24+1.42•1+0.24•1 = 5.4. The employment potential is 5.4 which means that the probability is over 75% that she will remain out of work for more than 4 months.

[12] One variant for taking into account the shift due to the constant is to go through the regression without the free term but with the full set of dummy variables for some parameter. In our case, the regression in column (5) of Table 1 was computed without the free term but with two dummy variables for gender (the first having the value of 1 for a woman and 0 for a man and the second, vice versa – 1 if a man and 0 if a woman). The coefficients obtained with the use of these dummy variables and additionally multiplied by 3 are presented in Table 2.

2.5.Additional Questions Recommended to Determine Employment Potential

As pointed out above, some characteristics determining the employment potential were not included in the statistical model because such data are either not in the base or are incomplete or are unfit for analytical processing.

The occupation of the unemployed is an example of an important factor that is hard to take into account. The data on occupation are contained in the personal card, but the principle of encoding occupations alphabetically makes this information impossible to use in a model. In the future, when new principles of encoding occupations are introduced or at least when changes are made in the personal registration card to allow for the use of an enlarged meaningful classification of occupations that factor can easily be integrated in the statistical model.

At this stage we have included in the questionnaire designed to determine the potential an additional question regarding the demand for the unemployed person's professional skills on the local labor market which simultaneously reflects the profession of the unemployed and the state of the local labor market. The profiling specialist must answer the question, "how easy is it to find work for an unemployed person who possesses a certain profession?" If unemployed persons of a given profession face approximately the same difficulties in finding employment as an average unemployed person, no additional points are awarded. If the unemployed of a given profession experiences maximum difficulties then 1 point is added to the employment potential index[13]. For a profession which is normally much in demand, on the contrary, 1 point is subtracted. If the difficulty of finding a job in a given profession is below (above) average one should subtract (add) 0.5 points. The choice of points assessment of the occupational affiliation factor at this stage is subjective which is inevitable given the lack of objective information[14].

In principle it is possible to include further additional questions to make the assessment of the employment potential more accurate. But the fact that the choice of weight coefficients for additionally included factors is arbitrary puts limits on their allowable number.

3.Method of Assessing Motivation for Job Seeking

To be effective the procedure of profiling unemployed citizens must take into account the motivation of the unemployed in terms of the risk of prolonged unemployment. To assess the motivation of the unemployed person on seeking a job two questionnaires have been developed. The first is a list of questions put by the profiling specialist to the unemployed and designed to determine the motivation to seek employment. The second questionnaire is a complement to the first and is a list of questions answered by the profiling specialist proceeding from the available data on the work of the unemployed person with vacancies during ten days since first registration with the employment service.

3.1.Questionnaire to Determine Motivation for Employment

This questionnaire is intended to determine the motivation of the employment service client to seek employment. The questions are put to the client by the worker of the employment center in the order indicated. Possible answers and corresponding points are determined for each question. The points are entered in the form for responses to determine motivation (See Form 1a) and are

[13] The employment potential index is formed in such a way that an increase of the index means smaller chances of finding a job.

[14] At the pilot stage we propose that if the employment center specialist disagrees with the assessment of the unemployed person's potential, he/she should adjust the assessment by indicating why he/ she disagrees.

then summed up.

Form 1a

¹	Questions	P o s		
1.	What is your purpose in coming to us?	The client says he ver to Question No.3) *2 points*		The answer has no reference to help in finding employment (over to *0 points*
2	Are you interested in seeking a job?	Yes *1 point*	No (over to Question No.) *0 points*	)Don't know (over to Question No.) 0 points
3	Why do you want to find a job? Why is it important for you to work?	Possible types of answers 1.Work is a source of income and guarantee of 2.Work confers social status and self-esteem 3.Work is necessary for communication and a 4.Work is a way to fulfill one's potential as a 5.Other The more categories are included in the answer nployment. Hard to say 0 points Refers to just one category 1 point Refers to 2-3 categories 3 points Refers to 4 categories or 4 points		
4	If a suitable offer comes along, are you ready to report to work within	Yes 2 points	Do not know 1 point	No 0 points
5	Are you ready to take a retraining or upgrading course if it helps you	Yes 2 points	Do not know 0 points	No 0 points
6	Do you agree that if you get no job offer for a long time it is better to accept any offer that comes along	Yes 2 points	Do not know 1 points	No 0 points
7	Do you think that many important events in your life were connected	Yes 2 points	Do not know 1 points	No 0 points
8	Can you afford not to work (for example, be a dependent of your	Yes	Do not know	No

		0 points	1 points	2 points
9	Would you be interested in looking for a job if you could draw	Yes	Do not know	No
		2 points	0 points	0 points

3.2.Behaviour Characteristics of High, Medium and Low Motivation for Employment

The second questionnaire contains characteristics of an employment center client's behavior from which his/ her motivation to job seeking can be deduced. For every characteristic three possible variants of behavior in work with vacancies are presented. The employment service officer can judge about them both from communicating with the client and from the entries in the "Job Search Plan"[15] The profiling specialist chooses the variant which, in his/ her opinion, corresponds to the client's behavior and gives a corresponding mark in the form (see Form 1b). The points are then added up.

Form 1b

¹	*Behaviour characteristic*	*Signs of high motivation for employment*	*Signs of medium motivation for employment*	*Signs of low motivation for*

[15] The registration form issued during first visit to an employment service registers the client's work with vacancies.

				employment
1	Nature of vacancies that interest the	*Considers vacancies that do not quite match his/ her skills or the* **2 points**	*Considers vacancies matching his/ her area of skills or the last wage* **1 point**	*Turns down reasonable vacancies* **0 points**
2	Wish to work	*Asks to be given as many* **2 points**	*Does not seek more* **1 point**	*Complains that he cannot visit* **0 points**
3	When he/ she	*Comes to the* **2 points**	*Comes to the employer,* **1 points**	*Comes to the employer on the* **0 points**
4	Attitude to	*Calls or telephones to* **2 points**	*Comes at the appointed* **1 points**	*Does not come or is late for* **0 points**
5	Interaction	*Cooperates with the* **2 points**	*Interacts with the* **1 points**	*Interacts reluctantly in a* **0 points**

At the pilot stage in the region we offer employment service workers a choice of questions from an additional list of questions below if they feel that this is necessary, or addition of some of their own questions or replace those that are in the questionnaires.

List of additional questions

Are you trying to find a job? How do you go about it?

Do you feel bored without work?

Will you try to start your own business if you don't find a job?

If a job in your field were only available in another city would you agree to move?

Are you ready to consider even those offers that do not quite match your skills?

Are you ready to consider offers with a smaller wage than you were paid in your last place of work?

If you won much money in a lottery would you prefer not to work for a while?

3.3.Deriving Summary Assessment of Motivation to Seek Employment on the Basis of Two Questionnaires

The summary assessment of motivation to seek employment is determined as a sum of points awarded for answers to the questions in the two above mentioned questionnaires. That is, the sum of points awarded for Form 1a is added to the sum of points in Form 1b. The total number of points ranges between 0 and 30.

It should be noted that, because we do not know the breakdown of the unemployed on the basis of motivation[16], the initial values of critical levels of motivation that cut off groups of unemployed people have been chosen arbitrarily and will be adjusted as information comes to hand regarding the breakdown of unemployed on the basis of motivation, on the one hand, and in line with the assessment of motivation by the employment center officer, on the other[17].

4.Procedure of Isolating Groups of Unemployed People on the Basis of Assessment of Prolonged Unemployment Risk

Ideally, the obtained assessments of the potential of employment and motivation for employment should be used to generate a single generalizing index, *risk of prolonged unemployment*, but lack of regular information on the motivation of the unemployed person prevents us from integrating the two dimensions of the risk of prolonged unemployment – employment potential and motivation – in a single index at the current stage because we did not know how to weigh these two indicators. So we have to resign ourselves to treating the risk of prolonged unemployment as a vector with two dimensions (employment potential and employment motivation). It has to be noted that the latter restricts the choice of the number of subgroups into which it would be reasonable to break down the total body of unemployed people because too many subgroups would be confusing.

[16] As noted above, regular data on the motivation of unemployed people are not collected.

[17] At the pilot stage we offer the employment center worker, if he/ she disagrees with the assessment of the motivation of the unemployed, to adjust the assessment stating the reason for disagreement. The distribution of points for responses in questionnaires may also be changed.

At this stage the Consultant proposes nine groups of unemployed people arrived at by combining three grades of assessment of employment potential (high, medium and low) and three grades for assessing the motivation for employment (high, low and zero). The boundaries of these groups may be variously drawn and the choice of critical values of employment potential and motivation is a management parameter that the regional Employment Department should adjust in accordance with the stringency of resource limitations.

Initially, we propose to choose the following critical values.

The potential of employment is considered to be low if the statistical model shows above a 75% probability of prolonged unemployment or, alternately, if the weighted sum of individual characteristics of the unemployed is above the critical value corresponding to the boundary of the upper quartile of the distribution of probability in the sub-sample used to evaluate the model[18]. Similarly, the employment potential is deemed to be high if the probability of prolonged unemployment in accordance with the statistical model is below 25% or, alternately, if the weighted sum of individual characteristics of the unemployed person is less than the critical value corresponding to the boundary of the bottom quartile of the distribution of probability in the sub-sample used to assess the model. The group in the middle is deemed to be a group with the medium potential for employment.

Motivation for employment: we propose that the motivation is high if the points aggregate in the two questionnaires is 20 points and higher; low if it is between 10 and 20 and zero if it is less than 10 points. Let it be repeated that testing in the regions will help to define more accurately the critical values for the motivation index.

Identification of nine groups of unemployed people completes the first stage of profiling.

5.Scheme of Selecting ALMPs Menu for Identified Groups of Unemployed

The second phase of profiling consists in selecting programs for each type of unemployed persons. Best practices of Russian specialists provide the methodological basis of the second phase of profiling. The methods of profiling used in the Chelyabinsk region and in the Republic of Karelia serve as a basis for determining the principles of directing certain groups of unemployed people to ALMPs. The methods are augmented by polling the opinions of employment workers in the pilot regions conducted as part of the project. A description of the methods of profiling and results of polling are presented in a separate material that will form part of the final report.

In the future it would be practicable at this stage to use a statistical model assessing the results of each type of program for various types of unemployed (as in Canada). But at the current stage data required for such an assessment are lacking.

A menu of the programs for each of the subgroups from which we recommend making a choice is presented in Table 3. As seen from the table, although we have identified nine groups we recommend concentrating efforts on the intermediate (gray) and not extreme groups.

[18] At the pilot stage we are trying out two variants. The second variant is preferable when the maximum predicted value of probability of prolonged unemployment is under 75%.

In our opinion, the “high potential, high motivation” group needs only to be informed about the state of the labor market and existing vacancies, and in the case of the “low potential, low motivation” it makes sense to just pay them unemployment benefits (after testing the validity of their claim to such a status). The test of the claim to unemployment status has been recommended for those who have no motivation to seek employment. For those with poor motivation programs to enhance motivation are recommended, while for those with a high potential programs of retraining and upgrading of job seeking skills are recommended.

Table 3.

Motivation	*High*	*Low*	*Absent*
Potential			
High	Pro-information on state atus r market and available		
Medium	n on state of labour ilable vacancies seeking tial for entrepreneurship ployment/ quotas		
Low	seeking ployment/ quotas		Just paying unemployment benefit

6. Automated Procedure

As pointed out above, the Consultant has prepared a mock-up of the automated profiling methods which represents a macros of the program formalizing the profiling procedure with the use of the EXCEL package. All the employment center worker has to do is introduce individual data of the unemployed in the electronic forms offered. On the basis of filled forms the potential, motivation and list of recommended ALMPs are automatically determined. The mock-up includes among other things.

- three electronic forms to be filled by the employment service officer (a form to determine potential, a form to determine motivation (Questionnaire 1), a form to determine motivation (Questionnaire 2)).
- automated procedure of assessing employment potential and assessing the motivation of a concrete unemployed person;
- possibility of taking into account the subjective opinion of the employment service officer through a change in assessment of potential and/ or motivation. At the testing stage an additional field is used designed to reflect why the employment service officer disagrees with the proposed ratings;
- The recommended choice of programs for a given unemployed person. At the testing stage the choice can be made both from the recommended list of programs or by enlarging and/ or altering the list. All the changes introduced by the employment service officer are recorded

7. Form of Determining Employment Potential that does not Require an Automated Procedure

In addition to the automated procedure we have prepared a form of determining the potential for employment (Form 2) that does not require automation. Form 2 is based on the questionnaire used in British Columbia, Canada. The points assessments used in the questionnaire are converted statistical estimates of the coefficients of the model discussed above. For ease of manual processing the coefficients have been rounded to whole numbers.

Questions 1-7 in Form 2 are filled by the unemployed persons. Thereafter Form 2 is filled by the employment center officer. Based on his experience she determines how competitive the unemployed person's professional skills are. If unemployed persons having this profession have approximately the same difficulty of finding employment as the average unemployed person, that is, the difficulty is average, no points are awarded. If unemployed persons with the given profession experience maximum difficulties, they are awarded three points. A profession which normally presents no difficulty, on the contrary, has 3 points subtracted. If the difficulty for a given profession is below (above) average one point has to be subtracted (added).

The resulting summary or adjusted number of points scored in answering questions in Forms 1a and 1b are entered in the "Degree of motivation" line in Form 2 filled by the unemployed person.

To determine the employment potential of a given unemployed person it is necessary to sum up the number of points for each question. The number of points for each response is within the service field, the far right gray-colored column. For answers to questions 1, 2 and 3 women are awarded the number of points in the left-hand gray column under the word "female" and men the number of points on the right-hand of the same column under the word "male." For other responses numbered 4-8 an equal number of points is awarded to everyone regardless of gender.

The summary number of points chosen by the unemployed in Form 2 may vary from 2 to 28. The number of points determines the potential of the given unemployed to find employment. A sum that equals 10 points or less means that the probability is at least 50% that the given unemployed person will find a job within four months after being registered. It means that among all the unemployed persons who have scored 10 points half will find work within the first four months. A sum that equals 16 or less means that the unemployed person will find work within 4 months with a probability of at least 25%. In other words, only a quarter of the unemployed who have scored 16 points will find work within 4 months.

Form 2. *Determining employment potential. Employment services in Voronezh Oblast (manual version)*

Choose one of the responses to each question.			
		Service field	
		female	*male*
1. Gender?	a. Female	5	
	b. Male		1
2. Education?	a. Elementary	0	0
	b. Secondary	0	0
	c. Secondary specialized	1	1
	d. Higher	1	2
3. Number of dependents in thea. None		0	0
	b. One	1	1
	c. Two	1	3
	d. Three	2	4
	e. Four plus	3	5
4. Age (years)?	a. Under 20	0	
	b. 20-24	1	
	c. 25-29	2	
	d. 30-34	3	
	e. 35-39	4	
	f. 40-44	5	
	g.45-49	6	
	h. over 50	7	
5. Total length of service, years?	a. None	5	
	b. Under three years	5	
	c. 3-9 years	4	
	d. 10-14 years	3	
	e. 15-24 years	2	
	f. 25-29 years	1	
	g. Over 30	0	
6. Length of service in the last placea. None		0	
	b. Under 5	0	
	c. 5-14	1	
	d. 15-29	2	
	e. Over 30	3	
7. Have you been dismissed as parta. Yes		4	
	b. No	0	
8. How easy is it for the unemployeda. Very hard		-3	
	b. Harder than average	-1	
	c. Average difficulty	0	
	d. Easier than average	1	
	e. Very easy	3	
Degree of motivation:		Total	

Conclusion:			
Date (date, month, year)			

8.Conclusion

In the light of international experience and the accepted practice of work with unemployed persons in Russia, the Consultant proposes a *mixed approach* to profiling, with the statistical model adjusted for the opinion of the employment service worker used in the first phase of profiling (to assess the risk of prolonged unemployment) and the opinion of the employment service worker totally relied on at the second phase (in selecting programs).

The merits of this procedure are obvious. Profiling draws on objective information about the unemployed person, the procedure is automated and user-friendly even if the skills of the local employment service workers are not high. The procedure is scientifically valid and permits concentrating the rendering of services to those unemployed who most need them.

The use of the experience of employment service workers compensates for some drawbacks of the approach that relies totally on the model because of failure to take fully into account the factors that may influence the length of the period out of work. The latter in turn is because at this stage some data that could have been taken into account in the model are inaccessible and some other factors do not lend themselves easily to formalization (motivation to work, the unemployed person's self-esteem).

At the initial stage in the use of this method in Russia the main problems stem from lack of unified formats of storing the data of personal registration cards and the codes used, which sometimes makes them unfit for analytical work. This calls for a large amount of preliminary additional work with data to bring them up to usable condition. Lack of resources and trained personnel in some regions may also impede the application of this approach.

Furthermore, an initial assessment of statistical model coefficients for the region is required as well as a periodic revision of the coefficients (once every 1-2 years). Regional academic centers could be instrumental in providing personnel for carrying out this work in a competent manner. If difficulties arise at the first stage it is possible (though not desirable) to confine the study to an economic region or a federal district.

However, it would be mistaken not to use the total body of personal information already being collected and to limit the classification of unemployed people to the more obvious, chiefly demographic factors.

The high potential and the skill of most workers of employment services leave no doubt that they are capable of learning the new methods as borne out by the experience of the regions that have started introducing similar methods.