

**FINAL REPORT**

**INDEPENDENT REVIEW**  
OF THE

**JOB NETWORK PROVIDER STAR RATINGS  
METHOD**

BY

**ACCESS ECONOMICS**

FOR THE

**STEERING COMMITTEE OF THE  
REVIEW**

**MARCH 2002**



While every effort has been made to ensure the accuracy of this document, the uncertain nature of economic data, forecasting and analysis means that Access Economics Pty Ltd is unable to make any warranties in relation to the information contained herein. Access Economics Pty Ltd, its employees and agents disclaim liability for any loss or damage which may arise as a consequence of any person relying on the information contained in this document.

---

---

**CONTENTS**

1.	Executive summary .....	1
2.	Introduction .....	6
2.1	Aims and Objectives of the Review.....	6
2.2	Conduct of the review.....	6
3.	Background to the review .....	8
3.1	The Job Network provider star ratings method.....	8
3.2	Concerns of Job Network members .....	13
3.3	Implications for the technical review .....	23
4.	Technical analysis .....	25
4.1	Analysis of technical modelling issues.....	25
4.2	Addition of new labour market and other variables .....	29
4.3	Secondary educational outcomes in the KPIs for Intensive Assistance.....	33
4.4	Assessment of NEIS.....	33
5.	Discussion and conclusions.....	34
5.1	Background .....	34
5.2	Implications of the technical analysis .....	34
5.3	Where to from here?.....	35
6.	Attachment A: Independent Review of the Job Network Provider Star Ratings Method: Terms of Reference .....	38
6.1	Background .....	38
6.2	Aims and Objectives of the Review.....	39
6.3	Timeliness and Reporting .....	40

---

---

## 1. EXECUTIVE SUMMARY

In December 2001 the Department of Employment and Workplace Relations (DEWR) commissioned Access Economics to undertake a technical review of the Job Network (JN) Provider Star Ratings Methodology. The main objective and scope of the review is to identify any potential for improvements in the control for differences in labour market conditions and job seeker characteristics.

A Steering Committee comprising representatives of DEWR and the chair of NESAs oversaw the review. The conduct of the analysis and the form and content of this report reflect their helpful guidance.

The review included extensive discussions with JN members in all capital cities and selected regional centres. We are grateful to all the JN members who contributed so substantially to our understanding of the issues.

In this final report of the review, we:

- ?? summarise the concerns expressed in the consultations that are relevant to our terms of reference;
- ?? report on the technical analysis undertaken;
- ?? make suggestions for refinement of the star ratings methodology; and
- ?? comment on other matters relevant to the terms of reference

### *Concerns raised in consultations*

JN members raised numerous concerns during the consultations. Some concerns reflected lack of understanding of the method or its application.

Others were directly relevant to our work. These included:

- ?? widespread concerns that the method did not adequately control for differences in labour market conditions in metropolitan, regional and remote parts of Australia;
- ?? views that boundaries of local labour markets were inappropriately defined in the model; and that other aspects of labour market conditions were inadequately represented;
- ?? concerns about inadequate representation of some job seeker characteristics, for example those with multiple disabilities;
- ?? concerns about the accuracy of the data and the treatment of some categories of job seeker.

### *Conduct of the technical analysis*

The technical analysis focussed on the modelling of Intensive Assistance (IA) and Job Search Training (JST). There was some consideration of relatively small amount of data available for the New Enterprise Incentive Scheme (NEIS), but no formal modelling of Job Matching.

The technical analysis had two main strands:

- ?? review of the statistical properties of the model; its functional form and use of data; and
- ?? preliminary exploration of possible expansions of the data set that would allow testing of concerns raised in the consultations.

In assessing the model, we took into account both a range of standard statistical tests, and also the impacts of possible changes on the estimated provider rankings and star ratings.

We undertook essentially the same analyses for Intensive Assistance and Job Search Training. The results are so similar that they allow us to draw conclusions that apply more or less equally to both service areas.

?? Given the similarity of the results for IA and JST - and the common approach and data sources - we believe that the results found for IA and JST would extend (with minor modifications) to Job Matching. Hence the overall conclusions of the review are likely to be applicable also to Job Matching.

Our objective was to review the adequacy of the current model and if necessary to suggest short term adjustments to mitigate any shortcomings - rather than to seek to establish new preferred specifications out of all the many permutations available. To achieve the latter may prove to be a substantial exercise in data collection, model formulation and analysis – beyond the resources and time available for the present study.

### *Main findings*

Main findings of the analysis are as follows:

#### *A) Review of statistical properties*

?? We were able, using the EViews econometric program, to replicate adequately the department's estimation results for IA and JST. The match was almost, but not quite exact. This was to be expected, given slight differences in rounding of data and in numerical methods within the software.

?? There is some evidence that more complex model specifications using the existing data set would improve the statistical fit of the models.

?? Changing the way that the provider effect is defined in the model has a relatively small impact on the estimated provider rankings or star ratings.

?? The estimated provider effects are themselves random variables, since they are derived from a statistical model. This means they are surrounded by a margin of uncertainty, that should be borne in mind in using the results from the model.

#### *B) Addition of new variables*

New variables are required to test the concerns about omitted labour market conditions raised during consultations. Only limited experimentation was possible during the course of the review.

?? It will be worth investigating the effects of varying the way the boundaries of local labour markets are defined.

Depending on the variable being defined, the “local labour market” is currently defined either as the Statistical Local Area where the job seeker is resident (which may be too small in many cases and too large in some others) – or else as the ABS Labour Market Region (which may be too large in some cases).

Exploring the issue would have required detailed adjustments to the data that could not be undertaken within the timeframe for this study

?? An experiment to assess the impact of including new variables relating to urban/regional differences in labour market characteristics suggested that they do have the potential to alter the estimated provider effects generated by the model. In the vast majority of cases the changes in star ratings were small. Nevertheless, **we believe that this points to an area where the model would benefit from immediate refinement.**

?? Further addition of new variables drawn from the Small Area Labour Force publication did not add to the explanatory power of the model for IA. We do not regard this experiment as particularly conclusive, since further research is needed into the definition of regional labour market boundaries and other issues.

- ?? Adding “dummy variables”<sup>1</sup> representing JN Labour Market Regions to the models for IA and JST has a large effect on the fit of the models and on the star ratings derived from them. This is to be expected, given the observed differences in average star ratings (and hence in estimated provider effects) across regions.

The dummy variables are picking up these systematic variations in average provider effects across regions. It is unclear how much of this in fact reflects unobserved differences in regional labour market conditions and job seeker characteristics, and how much genuine differences in provider effects.

Hence we do not advocate the use of such dummy variables in the model without further analysis of the underlying reasons for regional differences and the extent to which they can be represented directly, rather than through the use of dummy variables.

- ?? Including variables to represent Centrelink office effects has major implications for star ratings. As with the dummy variables for labour market regions, the dummy variables for Centrelink offices may be picking up a variety of effects (including differences in provider performance) besides those specifically relating to differences in Centrelink administrative practices etc. Therefore there is a need for further analysis of the reasons why the impacts are so substantial before any move to include these variables in the model.

Based on the limited experiments we were able to conduct in the time available, we draw the following conclusions relating to the addition of new variables:

- ✂ there is evidence that adding new variables to represent differences in regional labour markets would improve the statistical properties of the model. It is also likely to have some impact on the estimated provider effects and star ratings, tending to reduce regional differences;
- ✂ it appears that, even if such variables are added, there are still likely to be large estimated provider effects, unexplained by such variables; and
- ✂ it would be possible to make quick adjustments to the specification of the model that would improve the extent of control for differences in urban/regional labour market conditions. This would involve adding variables for: metropolitan/non-metropolitan location; residence in a Statistical Local Area with area over 2,000 hectares; and for the proportion of employment in different industries.

#### C) *Adjustment of the weight given to secondary educational outcomes in the KPIs for IA*

- ?? Reducing the weight on secondary educational outcomes in the KPIs for IA would affect the star ratings of 45% of JN members – mostly by 0.5 star. There is some evidence that providers in metropolitan areas would receive higher star ratings, while a very few non-metropolitan providers would lose 1 or 1.5 star. It did not appear that specialist providers would fare worse than other providers.

#### D) *Assessment of NEIS*

The analysis of NEIS differs from that of IA and JST because:

- ?? only a relatively small amount of data available is available and
- ?? there is no established set of models for NEIS for controlling for differences in labour market conditions and job seeker characteristics.

We tried applying to NEIS the model used for IA and JST. We found that:

- ?? the estimated models for NEIS outcomes have few significant variables. That is, outcomes are not significantly affected by labour market conditions and job seeker characteristics, as currently specified; and

<sup>1</sup> The “dummy variable” for a JN labour market region (e.g. the A.C.T.) takes the value “1” for job seekers that are clients of a JN member operating in that region. It takes the value “0” for all other job seekers. There is a different dummy variable for each JN labour market region. Adding them to the model controls for observed differences between regions in the mean probability of successful labour market outcomes.

- ?? the provider effects can be estimated using the raw data (i.e. with no corrections for labour market conditions and job seeker characteristics). For small numbers of providers with outcomes well above, or below, the average, the provider effects are significantly different from zero.

We conclude that:

- ✘ there may be sufficient information in the data to warrant assessing small numbers of NEIS providers as performing significantly above or below the average. However, application of the full star ratings methodology, as for IA and JST, is not justified, given the small amount of data on which to base the assessment of performance.

### *Overall assessment*

The use of a model such as that developed by DEWR to assess performance is a sound, leading-edge approach to performance measurement. The statistical model used to determine provider relativities has the advantages of objectivity, replicability, and reliance on data generated by the Job Network's operation. These are considerable advantages. There is no obvious replacement approach that can produce comparable national, comprehensive, objective assessments.

While conceptually sound and operationally reliable, the model for correcting for job seeker characteristics and labour market conditions is capable of refinement. We suggest that:

- ✘ the model be improved in the short run by adding a limited set of new variables controlling for urban and regional differences in labour market conditions, as described above; and
- ✘ over the medium term, there be a programme of systematic research to improve the data and the model in order to improve the identification of provider relativities. This research should focus on omitted potential variables that vary systematically across labour market regions, or with particular categories of JN provider.

There are inherent uncertainties about provider effects derived from a statistical model. These are magnified by the likelihood that some further revisions to the model may prove desirable in due course. This suggests that:

- ✘ consideration be given to reducing, at least temporarily, the number of star rating intervals by eliminating half-star ratings and using only whole numbers between 1 and 5. This would reduce the tendency of ratings to alter in response to small perturbations in the data or model;
- ✘ where possible other information should also be considered, in order to buttress star ratings that reflect rankings that are close to the cut-off values of the star distribution.

It would also be wise to use relevant additional information, as well as the star performance ratings, in assessing overall provider performance – particularly in critical business decisions.

- ✘ DEWR should consider using other relevant information besides the star performance ratings generated by the model to guide its business decisions.

We can identify three alternative sources of information that may help improve robustness of assessments of provider relativities. They are:

- ?? prices paid for JN services in different regions. These reflect, in part, a market assessment of the expected difficulty of achieving outcomes in different regions;

- ?? calibration of regional performance differences using information on comparative performance of nationally organised providers across regions; and
- ?? reducing the weight given to star ratings based on the model in favour of direct assessments of qualitative and other factors in provider performance, presently excluded from the assessment.

Each has potential advantages and disadvantages that should be considered before deciding to make use of the information.

*Access Economics*  
*March 2002*

---

---

## **2. INTRODUCTION**

In December 2001 the Department of Employment and Workplace Relations (DEWR) commissioned Access Economics to undertake a technical review of the Job Network Provider Star Ratings Methodology. The terms of reference for the review are at Attachment A.

### **2.1 AIMS AND OBJECTIVES OF THE REVIEW**

As set out in the terms of reference, the aims and objectives of the review are:

With a view to ensuring current best practice, the Review will assess the way in which the method used to rank Job Network members against current performance measures for each service accounts for differences in labour market conditions and job seeker characteristics. To inform potential future ratings development for Intensive Assistance in particular, the Review will examine the method used to combine employment and education outcomes to produce composite performance measures. The performance measures for all services are linked to the overall objectives of Job Network and the related contractual KPIs and therefore fall outside the scope of the Review. The relative weights of each performance measure are policy matters and so would also be outside the scope of the Review.

The specific scope and objective of the Review is to:

- ?? identify any potential for improvements in the control for differences in labour market conditions and job seeker characteristics, including substantive guidance on the implementation of any such refinements; and
- ?? examine the impact on relative performance scores of any change to the counting of Intensive Assistance secondary education outcomes to produce interim and final outcomes performance measures.

As part of the Review, consultations with NESAs and Job Network members will be conducted to identify areas for potential refinement of the ratings method and issues that have emerged within the industry associated with the use of the ratings system.....

In addition to producing reports, the Review team will provide observations to the Steering Committee (which incorporates representatives from DEWR and NESAs) on key issues emerging from the consultations related to the use of the Star Ratings system.

The review is timely, given the potential use of the star ratings as a factor in future purchasing decisions.

### **2.2 CONDUCT OF THE REVIEW**

The terms of reference required two parallel, interacting streams of work.

#### ***Consultations***

First, the review team<sup>2</sup> held consultations with Job Network members in every capital city and selected regional centres (Alice Springs, Townsville, Orange, Newcastle and Bendigo). These occurred between 15 and 23 January. Some 450 representatives of Job Network attended. There were also fourteen written submissions to the review.

The review team is grateful to DEWR and its regional offices for its efficient assistance in organising the venues and invitations, and to all the Job Network representatives who contributed to the meetings and submissions. The team gained many valuable insights that have helped shape the course of the study.

---

<sup>2</sup> Gavin Anderson & Company provided support to Access Economics in this phase of the study.

---

The issues arising during the consultations have been reported (in full and without attribution) to the Steering Committee. Those that are relevant to the technical review are summarised in section 3 of this report.

### *Technical analysis*

The technical analysis proceeded in parallel with the consultations. DEWR briefed Access Economics on the details of the methodology used in the regression analyses and the subsequent steps in the creation of provider rankings and star ratings. DEWR provided:

- ?? copies of research papers by the South Australian Centre for Economic Studies that formed the basis of the statistical models predicting job seeker outcomes;
- ?? lists of variables, regression results for each outcome equation, and definitions of the Key Performance Indicators (KPIs); and
- ?? data sets for the regressions for Intensive Assistance (IA), Job Search Training (JST) and the New Enterprise Investment Scheme (NEIS). These data sets included additional variables drawn from the Labour Force Survey, plus additional geographic indicators, developed by DEWR following discussion with the review team.

It quickly became apparent that undertaking a rigorous technical review in the short time available was a highly ambitious task. The data sets are very large. The data set for IA contains over 390,000 observations; that for JST over 87,000. Despite the best efforts of DEWR staff, there were delays in providing the data – largely reflecting the enormous task of ensuring quality for so many observations.

For the review team, the challenge was to establish an efficient method of handling such large data sets. It proved possible to solve the regression models quite rapidly on a personal computer, using the EViews econometric package. However, the associated manipulation of data and results required investment of time to establish efficient procedures.

Following discussion, the Steering Committee indicated priorities as follows:

1. review the regression models for IA outcomes, particularly the adjustments for labour market characteristics and job seeker characteristics;
2. explore the consequences for provider rankings and star ratings of altering the weight given to secondary education outcomes in the KPIs for IA;
3. review the regression models for JST and the adjustments for labour market characteristics and job seeker characteristics;
4. explore the potential for developing regression models to allow prediction of outcomes for NEIS. [The data set for NEIS is an order of magnitude smaller than that for JST and there are comparatively few final outcomes in the data set].

The Steering Committee indicated that review of the adjustments made in the regression models for Job Matching outcomes was of lower priority. In the event, no data were provided within the time frame of the review, and no modelling analysis was undertaken. However, we judge that the results of the analysis of IA and JST would be applicable also to Job Matching.

Consistent with these priorities, the review team began by analysing thoroughly the models and adjustments for IA. Having tested hypotheses and explored alternatives for IA, we then drew on the insights gained in designing a more limited set of key experiments using the data for JST and NEIS.

The analysis and findings were reported in detail to the Steering Committee, and are covered, with less technical detail, in this report.

### **3. BACKGROUND TO THE REVIEW**

This section has two roles. It provides a brief description of the Job Network provider star ratings method. It also summarises the concerns expressed in the public consultations that are relevant to the task facing the review team.

#### **3.1 THE JOB NETWORK PROVIDER STAR RATINGS METHOD**

##### *Role of the star ratings*

The Job Network (JN) ratings method ranks JN members according to their performance against a range of job seeker participation and outcomes measures. These measures are linked to the Key Performance Indicators (KPIs) and contractual conditions set out in the Employment Services Contract. The KPIs are listed in Table 1.

On the basis of their ranking relative to other service providers, each JN member is attributed a “Star” rating. Ratings range from “1 Star” to “5 Stars” with half Star increments. JN members are rated for each of the services they deliver in the regions in which they operate.

As part of JN’s competitive market framework, Star ratings are released publicly for each of the 19 JN regions with the aim of informing the individual job seeker’s choice of service provider and informing the wider public about the relative performance of JN members. Ratings for the 137 Employment Services Areas (ESAs) that make up the JN regions are released to JN members only.

**TABLE 1. JOB NETWORK MEMBER KEY PERFORMANCE INDICATORS**

<b>Service</b>	<b>Performance Measures</b>
Job Matching	<ul style="list-style-type: none"> <li><i>Job placements as a proportion of contracted number.</i></li> <li><i>Proportion of placements that are full-time jobs.</i></li> <li><i>Proportion of job placements taken up by: job seekers who have been unemployed for more than 6 months; and for more than 12 months.</i></li> <li><i>Proportion of job placements taken up by: people with disabilities, people from non-English speaking backgrounds, and Indigenous peoples.</i></li> </ul>
Job Search Training	<ul style="list-style-type: none"> <li><i>Proportion of job seekers placed in jobs within 3 months of their participation in Job Search Training.</i></li> <li><i>Proportion of job seekers for which a JST bonus payment was paid.</i></li> <li><i>Proportion of job seekers no longer receiving income support payments 3 months after their participation in Job Search Training.</i></li> <li><i>Proportion of job placements for Job Search Training participants taken up by: people with disabilities, people from non-English speaking backgrounds, and Indigenous peoples.</i></li> </ul>
Intensive Assistance	<ul style="list-style-type: none"> <li><i>Proportion of Intensive Assistance participants placed in jobs of a minimum 13 consecutive weeks' duration – and resulting in income support payment reductions of at least 70 percent.</i></li> <li><i>Proportion of Intensive Assistance participants placed in jobs of a minimum 26 consecutive weeks' duration – and resulting in income support payment reductions of at least 70 percent.</i></li> <li><i>Proportion of Intensive Assistance participants receiving job placements.</i></li> <li><i>Proportion of job placements for Intensive Assistance participants taken up by: people with disabilities, people from non-English speaking backgrounds, and Indigenous peoples.</i></li> </ul>
New Enterprise Incentive Scheme (NEIS)	<ul style="list-style-type: none"> <li><i>Number of job seekers commencing in NEIS as a proportion of contracted commencements.</i></li> <li><i>Proportion of NEIS participants no longer receiving income support 3 months after the end of programme participation.</i></li> <li><i>Proportion of NEIS participants withdrawing from NEIS and returning to income support.</i></li> </ul>

For all four services, only those job placements of sufficient duration to attract Job Network fees are counted for performance assessment purposes. Job Network members may place eligible job seekers into additional short-term temporary or casual jobs that, due to their brief duration, do not attract fees.

### ***Purpose of the JN Member ratings***

The ratings method was designed to allow for direct comparisons between JN members operating in different JN regions and ESAs. When assessing achievement against individual performance measures, regression modelling is applied to control for the potential impacts of differences in:

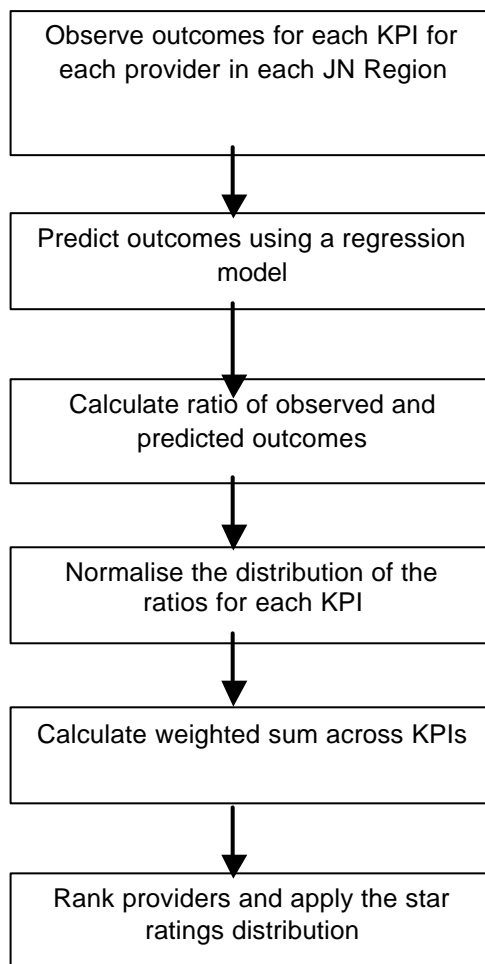
- ?? the labour market conditions under which JN members may operate: and
- ?? the characteristics of the job seekers they assist.

Ratings are also derived using an additional measure of JN members' success in assisting job seekers from three specific disadvantaged groups: indigenous people; people with disabilities; and people from non-English speaking backgrounds. This equity related measure, which reflects the particular emphasis the Government places on providing assistance to these disadvantaged job seekers, is not subject to control through regression.

Each performance measure is given a relative weight and then a combined ranking from all performance measures based on the weights is used to produce the "Star" rating.

*The methodology in more detail*

The methodology for calculating star ratings for a particular service (JM, IA, JST or NEIS) has the following steps.



We briefly detail each of these steps.

1. *Observe outcomes for each KPI for each provider in each JN Region*

?? The outcomes are those achieved by jobseekers relevant to the KPI. For IA, for example, the KPIs relate to:

achievement of a job placement (“Placement”);

placement in a job of a minimum 13 consecutive weeks’ duration – and resulting in income support payment reductions of at least 70 percent (“Interim Outcome”); and

placement in a job of a minimum 26 consecutive weeks’ duration – and resulting in income support payment reductions of at least 70 percent (“Final Outcome”).

?? For each outcome, calculate the number of positive outcomes achieved by each provider’s job seeker clients during the contract period. The calculation for each outcome includes only those job seekers that have been in assistance for a sufficient period of time (i.e. at least 26 weeks in the case of a Final Outcome).

2. *Predict outcomes using a regression model*

?? The regression model predicts the probability of a successful outcome for each job seeker, based on local labour market conditions and his/her characteristics. The explanatory variables in the model are:

?? for labour market conditions:

the local unemployment rate - in the Statistical Local Area (SLA) of residence

growth of employment in the local Labour Market Region

?? for job seeker characteristics:

gender

age

indigenous Australians

people with disabilities

people from non-English speaking backgrounds

sole parents

educational attainment

duration of unemployment

?? There is a separate model for each outcome, estimated using all the observed outcomes for all job seekers across Australia that have been in assistance for a sufficient period of time during the contract period.

?? For each JN provider in each Labour Market Region, use the regression estimates calculate the probability of a successful outcome for each of their job seeker clients, taking into account his/her characteristics and the local labour market conditions. Calculate the sum of these probabilities for each JN provider.

3. *Calculate the ratio of observed to predicted outcomes*

?? For each outcome, for each JN provider in each region, calculate the ratio of the number of positive outcomes observed (*at step 1*) to the total predicted number of positive outcomes (*at step 2*).

?? Note that this is equivalent to performing the same calculation using the KPIs themselves, rather than the numbers of outcomes. (This is because the KPIs are themselves ratios, each with a denominator that does not change. So it cancels out if we take ratios of observed to predicted.)

4. *Normalise the distributions of ratios for each KPI*

?? To allow the ratios for each KPI to be combined, they are first truncated (with any extreme values being reduced to a preset value – positive or negative as appropriate). The distributions of the ratios for each KPI are then normalised to have means and standard deviations both equal to 1.

5. *Calculate weighted sum across KPIs*

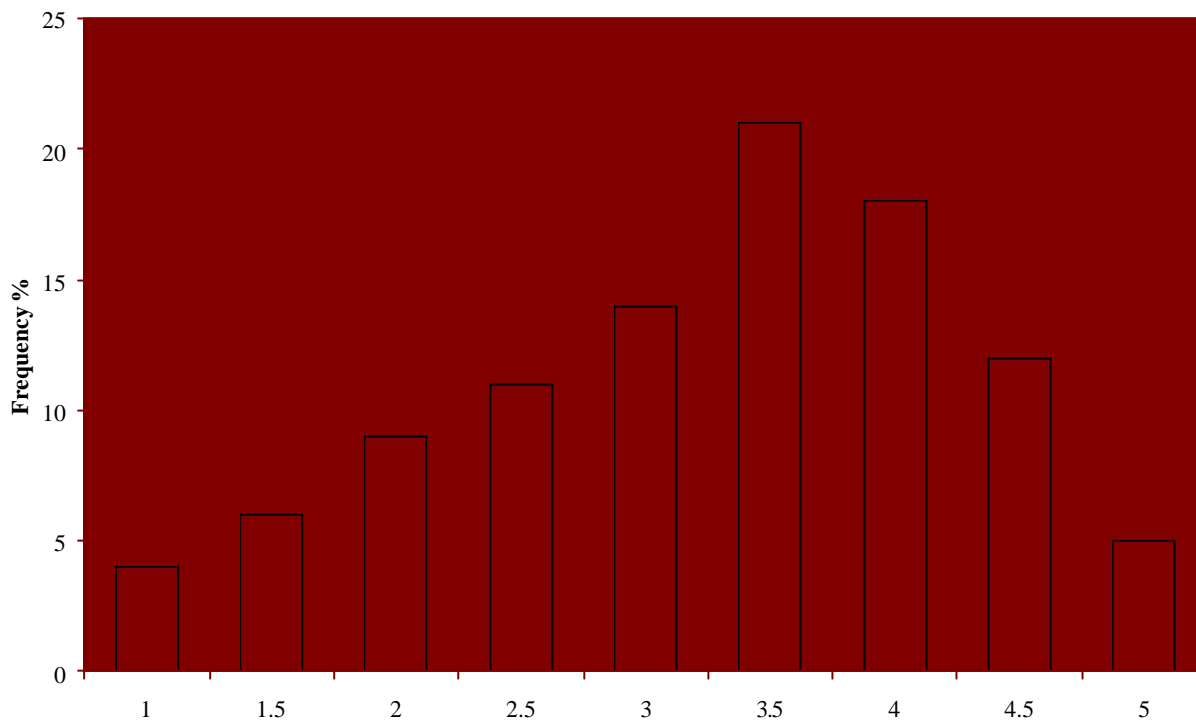
?? Use predetermined weights to calculate a weighted sum of the normalised ratios for each KPI, for each JN provider in each region.

6. *Rank providers and apply star ratings distribution*

?? The weighted sums (*from step 5*) are put in rank order. Stars ratings are allocated according to rank order, using a predetermined distribution of stars in the range +5.0 to 0, in half star intervals.

?? The distribution of stars is as shown in Figure 1.

**FIGURE 1. DISTRIBUTION OF STAR RATINGS (PERCENT OF JN MEMBERS)**



**3.2 CONCERNS OF JOB NETWORK MEMBERS**

The consultations held with Job Network members raised numerous issues concerning the adjustments for labour market conditions and job seeker characteristics that are central to the regression model. They also provided other suggestions that may be relevant to the terms of reference of the review.

All relevant concerns expressed by JN members are considered seriously. Some we were able to investigate during the course of our study. Others will need further analysis or testing, before their impact on the model can be determined.

A full report on the consultations is being provided separately to the Steering Committee. Here we summarise the issues and suggestions for change most relevant to the technical review. We also interpolate comments that interpret the points raised in modelling terms.

The questions raised in the consultations concern:

- ?? appropriateness of a national model
- ?? the logic of the outcomes
- ?? labour market conditions/regional urban bias
- ?? job seeker characteristics
- ?? quality adjustments
- ?? size of provider

- ?? achievement against contracted number of placements
- ?? specialist versus generalist suppliers
- ?? educational outcomes
- ?? data limitations
- ?? persons moving or finding employment

Not all are definitely relevant to this review. Some may reflect other aspects of the operation of Job network. In the following pages, we first summarise (*in italics*) the concerns raised in the consultations. We then briefly discuss (using an “arrow” bullet point) the possible modelling implications of each of the above issues.

### **Appropriateness of a national model**

Concerns raised in consultations:

*The issue is whether, given the complexities of and variations in regional conditions and the labour market, a national approach could, or should be undertaken in comparing performance.*

*The view was put that, given the complexities in accounting for regional variations in employment opportunities and in job seeker characteristics, it might not be possible to construct a robust national model. Perhaps a simpler system involving benchmarking, or “like-with-like” comparisons with a limited number of providers having similar characteristics (service profile, regional characteristics) would be more useful, and offer more accurate assessments.*

*Given that there is a limit to the extent to which the model could account for specific regional variations and other differences in labour market and job seeker characteristics was limited by the availability of the data – is there an appropriate trade-off between comprehensiveness and simplicity, and detailed coverage of the possible variables? Has lack of availability of data biased the model too much towards simplicity, meaning that it is unreliable as a guide to relative performance at a national level?*

Relevance to the technical review:

- ↘ In modelling terms, the question is: do regional differences, or other influences not accounted for in the model, mean that the results generated by the model are an unreliable guide to relative performance at a national level? This is the basic purpose of the technical review.

### **The logic of the outcomes is questioned**

Concerns raised in consultations:

*A common theme was that the outcomes from the model did not make sense. It was for example argued, principally but not exclusively by national providers, that the substantial differences in star ratings between and within regions did not reflect their assessments of what was happening on the ground.*

*How could the ratings be so different between Melbourne and Sydney, and between, for example, the Hunter region and South East Queensland? Given that similar management and service delivery systems were in place for national bodies represented in these different regions, the substantial differences in star ratings were cited as evidence that the model must be wrong. (In no case was it accepted that these differences might represent differences in the quality of the service offered by Job Network members).*

Relevance to the technical review:

- ↘ In modelling terms, the question is: are regional differences in labour market and job seeker characteristics adequately represented in the model? This again is central to the technical review.

---

---

## Labour market conditions/regional urban bias

Concerns raised in consultations:

*A dominant theme throughout the discussions was that labour market conditions were not reflected adequately in the regressions – and that as a consequence there was a bias in the results between urban and regional Australia. There are several facets to the argument.*

*Regional providers argued that the differences within and between regions in labour market characteristics were so great that it would not be possible to account for these differences – and that there was therefore an element of bias in the star ratings.*

Relevance to the technical review:

☞☞ At a gross level, it is possible to test for the existence of such differences by including dummy variables for different regions, or for metro/regional location. However, this does not provide clear evidence based on differences in labour market or job seeker characteristics. Any significant differences might still be caused by systematic differences in the quality of the service providers across regions or between city and country.

*It was argued that even if individual postcode data were to be used to control for differences in labour market characteristics, this would still not resolve the alleged bias in the star ratings as in regional and more remote areas of Australia there were substantial differences in labour market conditions between towns/regions within individual postcodes.*

Relevance to the technical review:

☞☞ This raises one aspect of a complex issue: the appropriate definition of the boundaries of local labour markets, and their relationship to variables currently used in the model - or potentially available should there be revisions to the data definitions. We explore this issue to some extent, but further work is required to explore its complexities.

*A common theme was that there were differences in labour market characteristics between the cities and regional (and remote) Australia because of structural change in the economy. Occupational and industry mix was an important determinant of outcomes. It was argued that structural change was producing growth in jobs in the cities that required high levels of education (for example in the IT industry), that these jobs were concentrated in the cities, and that the clients presenting to Job Network members (particularly for Intensive Assistance) did not have the skills to take up opportunities in such areas.*

Relevance to the technical review:

☞☞ This raises interrelated issues of (A) the rate of structural change in job characteristics in local labour markets; and (B) the flow of new jobs and their characteristics (industry, occupation etc.) relative to the numbers and skills of Job Network clients. Neither issue would seem to be directly reflected in the labour market variables currently included in the model. The empirical significance is yet to be established.

*The size of the local labour market and the number of employers was important, in that it affected choice, diversity and the number of jobs becoming available in any one period.*

Relevance to the technical review:

☞☞ Accepting that this issue may be important, it is not represented by the labour market variables currently included in the model.

*In some areas, it might be the case that new investment and growth could be assumed to generate new job opportunities, but in reality the jobs might not be available for locals. For example, a highway or other*

---

---

*infrastructure project might result in job opportunities for a region being assessed as being propitious. But in reality the company with the contract for this work would tend to bring in its own labour, with the result that few, if any, jobs were available for locals.*

Relevance to the technical review:

- ✂✂ While this may appear an issue of detail, it could be important in some areas where there are large construction projects. An analogous issue arises in the case of fly-in fly-out mining operations. It may affect the accuracy of variables purporting to measure growth in, or number of jobs available to, local residents.

*The trend to casual and part-time work was not reflected properly in the model. The suggestion was that this trend was more apparent in regional than urban Australia.*

Relevance to the technical review:

- ✂✂ This is potentially important, in that a higher proportion of casual and part-time employment may make it more difficult to achieve long term job placements or required degrees of reduction in benefit. It may be possible to obtain direct estimates of the proportion of casual and part-time jobs in local labour markets, or else to use proxies based on industry mix.

*Another source of bias was alleged to be the differences in distribution of category A and category B clients between regional and urban areas. It was argued that there was a close correlation between star ratings and the distribution of these two categories.*

- ✂✂ If there is in fact an issue of concern, the implications of this for the modelling are not immediately clear. It could relate to the detailed definition of variables representing job seeker characteristics, or possibly to omission of some variable with which the effect is correlated.

*It was suggested also that there might be a non-linear relationship between the local unemployment rate and the ease of achieving outcomes. Where unemployment rates were very high, it was difficult to achieve outcomes because there were few job opportunities per job seeker. Where unemployment rates were very low, those who presented as clients usually had severe problems that meant that outcomes were difficult to achieve.*

Relevance to the technical review:

- ✂✂ It would be possible to test this hypothesis by modifying the form of the existing regression model to include a non-linear relationship between probability of an outcome and the local unemployment rate.

*The model could not account properly for seasonality in the labour market, nor for differences in, for example, agricultural output. Whether there was a good or bad agricultural season had a substantial impact on employment prospects in regional Australia, but this was not reflected in the model.*

Relevance to the technical review:

- ✂✂ This is relevant, in that a higher proportion of seasonal employment may make it more difficult to achieve long term job placements or required degrees of reduction in benefit. It may be possible to obtain direct estimates of the proportion of such jobs in local labour markets, based on the observed quarterly variation in job numbers, or else to use proxy variables based on industry mix.

*It was argued in Darwin that whereas the wet season had a major impact on the ability of Job Network members to service clients in more remote areas, this was not reflected in the model.*

Relevance to the technical review:

- ✂✂ This is a possible argument in favour of a regional (or climatic) dummy variable. It would also be possible to attempt to identify directly those clients likely to be affected.

---

---

*Motivational factors were also seen as important. In some centres people coming into Job Network centres were motivated to avoid work and remain on welfare. This was not counted in the regressions, but had a major impact on the ability of Job Network members to achieve outcomes for such clients – who were skilful in exploiting the system.*

Relevance to the technical review:

☞☞ Motivational and other long-term cultural factors that affect the ability of providers to achieve outcomes might be controlled for by including a variable such as the local participation rate in the model.

*Transport issues were also seen as a source of regional/urban bias in the model. However, even within cities there were problems. Transport issues were perceived as relating to access, cost and distance. Access and cost tended to be worse in regional than urban Australia. The current application of the 90 minute rule was inappropriate in regional Australia as travel times to work could be much longer, and/or distance could be covered more quickly – implying higher transport costs per minute. Public transport was not as accessible to clients in regional centres as it was in the cities. Indigenous Australians were at a particular disadvantage. The more remote the area, the worse the problem.*

*The issue could be seen as applying at the level of the individual jobseeker (e.g. no car and/or driving licence), or as a characteristic of local labour markets (job dispersion; availability of public transport).*

Relevance to the technical review:

☞☞ This appears to be a significant issue, that does not feature explicitly in the current model. It is a question of measuring the number of jobs reasonably accessible by a job seeker – this means (A) defining appropriately the boundaries of the local labour market and (B) controlling adequately for job seeker characteristics that affect mobility.

*The methodology should group similarly placed regions together for the purposes of the regressions. The argument was that northern WA, northern Queensland and the Northern Territory had more in common with each other than they did with the rest of Australia – especially as regards remoteness and also in terms of the higher proportion of their populations who were Aboriginal. They should therefore be grouped together.*

Relevance to the technical review:

☞☞ If it were judged that a single national regression model were unable to control adequately for the differences, then applying separate models to regional subsets of providers would be a possible option.

*Some (e.g. in Queensland) suggested that local government data should be incorporated in the regressions. Others were however of the view that there was too much variability in this data between states for it to be useful. Other data possibilities mentioned included: dwelling rentals and house prices; use of NES apprenticeship targets as a direct measure of local demand for labour.*

Relevance to the technical review:

☞☞ Suggestions for completely new variables need to be considered with an open mind, balancing costs against potential contribution to the model. The review team also considered the question of obtaining new data, but had neither the time nor resources to explore this issue to any extent. The report does make suggestions for further work on this aspect.

*Another concern was that the methodology tended to penalise providers that had opted to provide a distributed service through a network of small offices, compared to those that had elected only to serve larger regional centres. This was partly a question of the difficulty of achieving outcomes in small labour markets.*

Relevance to the technical review:

- ✂✂ This concern may not be fully captured in variables representing the size of local labour markets. It is possible that variables representing provider configuration may provide relevant additional information on the characteristics of local labour markets.

### **Job seeker characteristics**

Concerns raised in consultations:

*While it was recognized that differences in job seeker characteristics were controlled for in the regressions, there is a view that these differences are not measured in the right way. Some felt that job seeker characteristics were more likely to be mis-specified in regional than in urban Australia.*

*Inaccurate job seeker characteristics, reflected in the responses to the Job Seeker Characteristics Index (JSCI) applied by Centrelink, were said to be particularly bad in regions where the proportion of indigenous clients was high.*

*It was said there were substantial differences between indigenous persons living in a remote part of Australia (where they were more likely to be living a traditional lifestyle) and those living in one of the main cities. There was a common view that there needed to be more differentiation of indigenous clients according to their English language proficiency. The extent of misclassification was said to be worse in regional than in urban Australia.*

Relevance to the technical review:

- ✂✂ The hypotheses here are that the characterisation of indigenous job seekers is either (A) too simple in the JSCI and/or in the derived variables used in the model; or (B) that errors in classification are greater in regional Australia, or (C) that there are systematic difference between local labour markets in the ease or otherwise of securing employment for indigenous clients. The issues here are too complex for substantial progress to have been achieved within the limits of this technical review. We note that job seekers may differ on other measures of disadvantage, besides their indigenous status.
- ✂✂ Note also the implied claim that there may be systematic differences in outcomes as a result of differences in behaviour by Centrelink offices.

*The operations of the Aboriginal Community Development Employment Program (CDEP) needed to be included in the regressions. One argument was that by putting people into training, for which there was allegedly no prospect of gaining sustainable employment, CDEP made measured unemployment among Aboriginal people in a region look better than it was in reality. Another was that if a jobseeker was placed into a CDEP funded job, then it did not count as an outcome under some KPIs. Hence the local prevalence of CDEP funding was a factor that needed to be included in the model.*

*Similar comments were made about placements in other Commonwealth training and group apprenticeship programs. Variations in their incidence could differentially affect the proportion of positive outcomes achievable.*

*A further issue raised in the consultations was that the difficulty of achieving a given proportionate reduction in benefit depended on the level of benefit being received by the job seeker. Some for example had larger supplements for dependents*

Relevance to the technical review:

- ✂✂ The issues are that the uneven geographic incidence of other Commonwealth programs (A) may affect the likelihood of positive outcomes being achieved: a factor outside the control of Job Network members, and/or (B) may bias other measures of labour market characteristics, notably the recorded unemployment rate. There is also the question of the level of benefit being paid to a job seeker.
- ✂✂ The issues overall seem worth attention, but their practical significance has yet to be assessed. New data would be required.

---

Concerns were also expressed that ability to achieve on KPIs that measured the proportion of indigenous or disadvantaged jobseekers placed depended on the number that were referred by Centrelink to the provider. This was seen as often being out of the control of the Job Network member.

Relevance to the technical review:

✂✂ This raises an issue whether KPIs relating to achievement of outcomes for disadvantaged clients should be adjusted for differences in labour market characteristics. This is outside the scope of the technical review .

It was also suggested that it was more difficult to place NESB clients where they formed a comparatively small proportion of the local population. This was because other support networks were less well developed, and there were fewer NESB employers who tended to be more willing to employ them.

Relevance to the technical review:

✂✂ This suggests that the size of the local NESB population (and possibly its ethnic composition) may influence labour market characteristics. We consider this to some extent in the course of the technical review .

## Quality adjustments

Concerns raised in consultations:

The need for some adjustment to the quantitative process to account for the quality of the service delivered was a key theme.

One way this was expressed was that the star system measured outcomes for only 20 to 30 percent of jobseekers. No weight was given to achievement (or lack of it) in relation to the other 70-80 percent of clients.

A common view was that reliance on the star system alone to determine contract rollover was inappropriate. Some assessment needed to be made of the quality of the service being offered. There was however no consensus as to the best way to do this. One suggestion was that around 80 per cent of the “mark” should be done via the star system, with around 20 per cent accounted for by an assessment of the quality of the product being offered.

Some suggested that this could be done by contract managers. Others suggested that DEWR’s audit reports could be used for providing quality adjustments, as could client satisfaction survey - perhaps having a quality assessor appointed for providing all such quality assessments within a region.

Relevance to the technical review:

✂✂ While this is an important set of issues, it is not directly within the scope of the technical review.

## Size of provider

Concerns raised in consultations:

It was argued that small providers were both advantaged and disadvantaged vis a vis larger providers. No clear consensus emerged.

Some argued that they were disadvantaged by being the single provider in an ESA, compared with multiple providers.

Relevance to the technical review:

✂✂ It is not apparent that this raises any issue relevant to the technical review.

---

---

*The comment was made that the price DEWR had to pay for outcomes and the fees for services, where no outcome-based contracts existed, should provide a direct market indicator of the difficulty of achieving outcomes in particular regions and service categories.*

Relevance to the technical review:

✂✂ The suggestion is that the price DEWR has to pay in an area to secure services may be a direct market indicator of the difficulty of securing a labour market outcome. This may warrant exploration in the model.

### **Achievement against contracted number of placements**

Concerns raised in consultations:

*There were frequent comments that, particularly in Job Matching, providers were penalised for achieving less than the contracted number of placements. While there was acceptance that in some cases this reflected unrealistic tenders, in other cases it resulted from adjustments to tendered numbers by DEWR, that had turned out inappropriate given the subsequent evolution of the labour market.*

Relevance to the technical review:

✂✂ Since Job Match was a lower priority for the technical review and the relevant KPI is not currently adjusted for labour market conditions, we have not investigated this further. However, prima facie, the number of Job Match places contracted by DEWR relative to the number of jobs to be filled in a local labour market, would appear to be a variable (A) not under the control of individual Job Network members and (B) likely to affect the extent of performance against the KPI. To the extent that this reflects a policy decision, it is outside the scope of the review.

### **Specialist versus generalist suppliers**

Concerns raised in consultations:

*There was a view that specialist providers were advantaged vis a vis general suppliers as they addressed the needs of those who were likely to be more likely to be skilled and/or motivated. A view was that this should be tested specifically in the regressions.*

Relevance to the technical review:

✂✂ While this could be tested in the regressions, it is unclear that it is a factor that requires controlling for in the model. Specialist providers may on the whole be more efficient than the general suppliers.

### **Educational outcomes**

Concerns raised in consultations:

*There was a strongly held view that there should be no short-term changes in the education weights. There had already been too many changes, and more would represent changing the goal posts mid game.*

*A general view was that educational outcomes were fundamentally valuable – whether as re-skilling, or in teaching basic employment necessities (such as language skills). Hence the KPIs should continue to recognise the importance of such outcomes. On balance it appeared there was majority support for no change to the current weighting of education in the KPIs.*

*It was wrong to believe that educational outcomes were generally easier to achieve than job outcomes. Motivating clients to attend, and to keep attending for a semester, was not easy in many cases.*

---

---

*It was recognised that in some instances, educational outcomes were being sought in cases where it was inappropriate so as to bolster star ratings. This was best dealt with by administrative sanctions rather than by changing the star ratings methodology.*

*There was a view that rural areas were disadvantaged vis a vis the cities in the provision of education outcomes. In addition to fewer courses being available, it was more difficult in regional areas to place people in relevant subsequent courses – as these tended to be cancelled at short notice given the lower numbers of students.*

*There was also a view that some providers were advantaged because of their links with education providers.*

Relevance to the technical review:

*✍✍ This is not relevant to the control for labour market conditions and job seeker characteristics. However, the Steering Committee asked that we investigate the consequences of removing secondary educational outcomes for the regression outcomes and the patterns of star ratings. We have investigated on this issue.*

### **Data limitations**

Concerns raised in consultations:

*Not necessarily reflecting fault, serious reservations were expressed about Centrelink data. If there could be no confidence in the data, there could be no confidence in the results.*

*It was noted that it was not Centrelink's main function to fill in the JSCI accurately – even if clients gave them accurate information. They were in any case under-resourced for the job they needed to do. There were also strong views that differing practices across Centrelink offices (in relation to data, classifications, referrals, and other administration) could affect the probability of achieving outcomes, independently of other factors controlled in the model.*

*It was suggested that there be some study made of the accuracy of Centrelink data.*

Relevance to the technical review:

*✍✍ While such concerns may be valid at the level of individual job seekers, they are of concern for the regression model only if there are systematic variations across areas or classes of job seeker. It is possible justification for including dummy variables to reflect the possible differential influence of individual Centrelink offices on the outcomes obtained by Job Network members. However, these are rather a blunt instrument for controlling for this effect.*

*There were also several suggestions that a greater range of Centrelink data be used in the regression models - e.g. occupation; access to transport; reason for exit. Historical analysis of Centrelink data (mapping postcodes of residence into postcodes of employers) would provide an objective way of defining local labour market areas.*

Relevance to the technical review:

*✍✍ It was not possible to explore the possibilities within this study. Prima facie it seems sensible to use fully the Centrelink data available. However, this needs to be balanced against known limitations in the quality of the dataset..*

### **Persons moving or finding employment**

Concerns raised in consultations:

*There was a view that the system did not account properly for those who moved residence, or had found employment on their own initiative – possibly helped by the Job Network member (e.g. as a result of Job Search Training). This was not reflected in the data.*

Relevance to the technical review:

~~///~~ Such issues may be relevant, but require further investigation to establish their significance. However, data were not available to allow us to explore them in the present study.

---

### 3.3 IMPLICATIONS FOR THE TECHNICAL REVIEW

#### *Assessing the methodology*

Derivation of star ratings involves two key steps:

- ?? comparison of actual and predicted outcomes obtained by each service provider, using a regression model of job seeker performance that controls for local labour market conditions and job seeker characteristics; and
- ?? combining the ratios of actual and predicted outcomes into a single index for each service area; using it to rank service providers across the country; and assigning star ratings according to the rank.

The technical analysis focuses on the adequacy and accuracy of the regression model, that predicts job seeker outcomes and furnishes an estimate of the “provider effects” (i.e. the additional contribution to the achievement of outcomes made by each service provider).

This naturally turns attention towards the statistical properties of the model; its functional form; and the properties of the difference between actual and predicted outcomes that it generates for each service provider.

However, for practical purposes we are also interested in the accuracy and robustness of the star ratings that result from the application of the model. To what extent would changes in the model to improve its statistical properties also cause changes in star ratings? What margin of uncertainty surrounds the star ratings, bearing in mind that they are estimates based on results from a statistical model?

#### *Assessing the concerns of Job Network members*

Job Network member’s concerns partly reflect lack of understanding of the details of the method of deriving star ratings; partly some scepticism about the results; and partly some more specific concerns about the adequacy of the model’s specification and data.

If there is an overriding concern, it probably relates to the perceived inadequacy of the control for labour market conditions. However, there are concerns about the quality of the data for individual job seekers, and the detail of control for job seeker characteristics. There are also issues about the data period that should be used in the assessment, and the treatment of some types of job seeker (e.g. those that move residence).

The concerns about treatment of labour market conditions are most evident in the frequent claim that the difficulties of achieving outcomes in thin labour markets in regional and remote Australia are not adequately represented in the model. However, there are also concerns about apparently odd differences between star ratings obtained in the various capital cities.

Testing these concerns involves several steps:

#### **reviewing the definitions of local labour market boundaries used in the model**

- ?? At present the unemployment rate is that in the statistical local area<sup>3</sup> (SLA) best matching the postcode of residence. Growth in employment is that in the relevant ABS Labour Market Region.

---

<sup>3</sup> Statistical Local Areas are a concept similar to Local Government Areas, used by the Australian Bureau of Statistics in producing nationally comparable regional statistics.

- ?? There are difficulties in accurately measuring many labour market variables derived from the Labour Force Survey for smaller areas than the Labour Market Regions. However, employment, unemployment and population are available for SLAs. Further detail is also available from the population census and other sources. Hence it is possible to review the boundaries of local labour markets used to define the variables in the dataset, to ensure that they most closely match the extent of employment opportunities available to jobseekers.

**testing the inclusion of relevant new variables that can readily be derived from the existing data set**

- ?? For example, it is possible to test concerns that the probability of an outcome varies non-linearly with the local unemployment rate; or that the ease with which a non-english speaking (NESB) person can obtain an outcome varies according to the size of the local NESB population<sup>4</sup>.

**obtaining new data for the variables necessary to test other hypotheses raised by JN members**

- ?? These could be based on the Labour Force Survey, on the Population Census or on other sources. In each case there is a need to weigh the cost and time to obtain and test the new variables, against the accuracy and relevance of the new variables, and their potential to materially alter the model and star ratings.

Similar considerations arise in relation to the control for job seeker characteristics. However, here the main focus is on the extent and accuracy of data available from the JSCI and the Centrelink database – as well as the definition of the variables included in the model.

---

<sup>4</sup> This can only be done approximately using the existing data set. It requires the (probably reasonable) assumption that the number of NESB jobseekers in an SLA reflects the size of the NESB population.

---

---

## 4. TECHNICAL ANALYSIS

The technical analysis followed the two tracks foreshadowed in the previous section:

- review of the statistical properties of the model; its functional form and use of data; and
- preliminary exploration of possible expansions of the data set that would allow testing of concerns raised in the consultations.

In assessing the model, we took into account both a range of standard statistical tests, and also the impacts of possible changes on the estimated provider rankings and star ratings.

In the time available we have not sought to derive the most perfect models attainable with the available data – rather to review the adequacy of the existing approach and formulation, and if necessary make suggestions for short term refinement. Given the difficulties of expanding the data set, we were able to make only some initial steps in exploring the possible implications. The results of this endeavour nevertheless provided useful insights that are reflected in the suggestions for further research.

We undertook essentially the same analyses for Intensive Assistance and Job Search Training. The results are so similar that they allow us to draw conclusions that apply more or less equally to both service areas. There are some differences of detail, caused by the different durations of assistance and KPIs.

The analysis of IA took place before that for JST. We therefore performed more experiments on the model for IA than on that for JST. Given the similarity of the other results, we were able to focus the effort on JST more sharply.

Given the similarity of the results for IA and JST - and the common approach and data source - we believe that the results found for IA and JST would extend (with minor modifications) to Job Matching.

The analysis of NEIS, by contrast, differs from that for IA and JST. Only a small amount of data is available – especially for the final outcome. Moreover, there is no established set of models for NEIS outcomes like that for IA and JST.

We do not attempt to present here the technical detail of the analysis. It has been provided separately to the Steering Committee. Rather we focus on broad findings and some basic issues that arose in the course of the analysis.

Our overall assessment is that the basic method of estimating JN member performance is sound, but is capable of refinement. The refinements potentially relate to:

- the technical form of the model;
- the interpretation of the estimated provider rankings and star ratings; and
- possible expansions and refinements of the data set, particularly relating to the representation of local labour market conditions.

The remainder of this section presents and discusses the findings of the technical analysis. The next section focuses on possible avenues for refinement of the present method.

### 4.1 ANALYSIS OF TECHNICAL MODELLING ISSUES

#### *Replicating DEWR results*

Because of the size of the data sets involved and the time it took to verify them, there were some initial delays in accessing the data. Once these issues were resolved:

- we were able, using the EViews econometric program, to replicate adequately the department's estimation results for IA and JST. The estimated star ratings were the same. The statistical match in the estimation

results was almost, but not quite exact. This was to be expected, given slight differences in rounding of data and in numerical methods within the software.

The data for the models consists of outcomes and other variables representing job seeker characteristics and local labour market conditions for each job seeker placed in assistance. The data include commencements from the beginning of the contract period until August 2001. Job seekers that have had insufficient time in assistance to qualify for an outcome are excluded from that regression.

For IA there are over 390,000 observations in the data set; for JST over 87,000. For NEIS there are up to 8,700 (but many of these do not have a final outcome).

**TABLE 2. OUTCOMES EXPLAINED BY THE REGRESSION MODELS**

INTENSIVE ASSISTANCE		JOB SEARCH TRAINING	
PLACED	job seeker commencement results in a Job Matching paid placement	PLACED	job seeker commencement results in a Job Matching paid placement
INTERIM	job seeker is placed in a job of a minimum of 13 weeks' duration – and resulting in income support payment reductions of at least 79 percent	JSTBONUS	job seeker is placed in a job of a minimum of 13 weeks' duration – and resulting in income support payment reductions of at least 79 percent
FINAL	job seeker is placed in a job of a minimum of 26 weeks' duration – and resulting in income support payment reductions of at least 79 percent	OFFB	job seeker is placed in a job of a minimum of 26 weeks' duration – and resulting in income support payment reductions of at least 79 percent

The outcomes for IA and JST, explained by the models are listed in Table 2. The dependent variables in the models (i.e. the outcomes explained) are binary in form (i.e. they take the value 1 if the job seeker is successful in achieving an outcome; 0 if unsuccessful). This means that standard regression analysis is inappropriate. Instead logit or probit models are most commonly used. DEWR uses the logit form.

The explanatory variables used in the models are listed in Table 3. DEWR uses the same variables in all regressions.

**TABLE 3. EXPLANATORY VARIABLES USED IN THE MODEL**

AGE	Age of individual
AGE2	Age squared <sup>2</sup>
ALST	= 1 if on assistance at the start of the sample period, = 0 otherwise
DAYSASS	Days of assistance provided, computed as the minimum of (365 or the number of days between the commencement date and 31 August 2001)
DIS1	Disability Dummy Variable
DIS2	Disability Dummy Variable
DIS3	Disability Dummy Variable
DIS4	Disability Dummy Variable
EDU1	= 1 if highest level of educational achievement is not volunteered, a special school, completed primary school, completed less than year 10, a special unit in a school or did not attend school, = 0 otherwise
EDU2	= 1 if highest level of educational achievement is Year 10 completed or Year 11 completed, = 0 otherwise
EDU3	= 1 if highest level of educational achievement is Secondary School completed, = 0 otherwise
EDU4	=1 if highest level of educational achievement is Trade/TAFE qualification or Associate Diploma, = 0 otherwise
EDU5	=1 if highest level of educational achievement is Degree or Post-Graduate Degree, = 0 otherwise
EG2	Regional employment growth (ABS Labour Market Region)
INDIG	= 1 if Aboriginal or Torres Strait Islander, = 0 otherwise
MALE	=1 if client is male, = 0 otherwise
NESB1	NESB Dummy Variable (according to country of origin)
NESB2	NESB Dummy Variable (according to country of origin)
NESB3	NESB Dummy Variable (according to country of origin)
SOLO	=1 if Sole Parent, = 0 otherwise
UE2	Local unemployment rate (based on client SLA of residence)
UEDUR	Number of months unemployed at start of assistance

### *More elaborate models using the existing data*

☞ There is some evidence that more complex model specifications using the existing data set would improve the statistical fit of the models.

This is indicated by some statistical tests, and also by an experiment in which we included in the equations several plausible compound variables<sup>5</sup> formed from those already in the model. Together these improved the fit of the model. They also cause some change in provider rankings and hence in star ratings – though in almost all cases, by no more than half a star. The impact of making these changes is larger for IA than for JST.

In a separate experiment, we tested the hypothesis that the probability of a successful outcome varies non-linearly with the local unemployment rate<sup>6</sup>. The data did not support the hypothesis for IA. The experiment was not repeated for JST.

### *Changing the way the provider effect is defined*

☞ Changing the way that the provider effect is defined in the model has a relatively small impact on the estimated provider rankings or star ratings.

The data allocate each job seeker to a JN provider in a particular JN region. A provider effect is derived separately for each of these region/provider combinations.

Provider effects are currently defined as the ratios of actual to predicted outcomes. Alternatives are to define the effect as: (A) the difference between actual and predicted outcomes; or (B) a random effect, adding to the existing error term in the equation. They each imply a slightly different distribution of impacts of provider effects across job seekers – a choice between them needs to be made on conceptual grounds.

It turns out, though, that altering the specification has a relatively small impact on provider rankings or star ratings.

### *The provider effects are themselves random variables*

Analysis of the provider effects highlights the fact that they are themselves random variables, each with its own margin of uncertainty. This is because they are derived from a statistical model, rather than by direct observation.

Figure 2 shows the distribution of provider effects for IA, as derived from the model, using the ratio method, and after calculating the weighted average across KPIs. This is prior to applying the distribution of star ratings determined by DEWR. These also are shown in the figure as a series of large dots<sup>7</sup>.

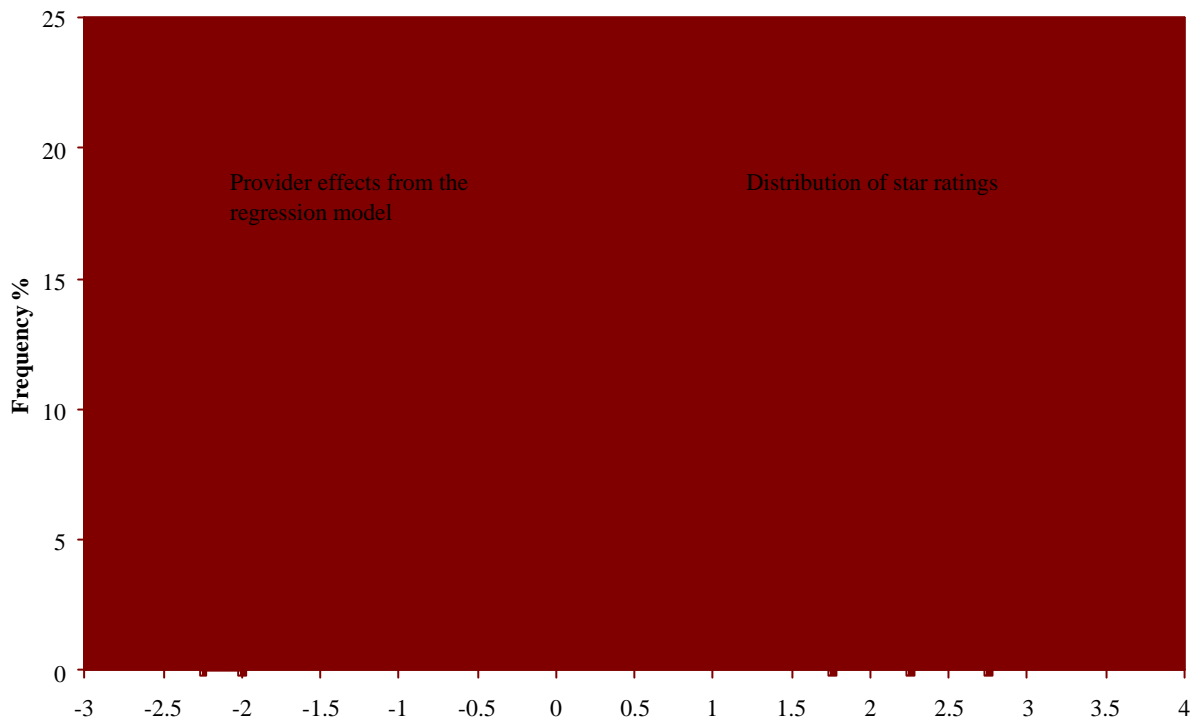
The provider effects, as estimated, are more tightly bunched than the star distribution. This means that small perturbations in the estimated provider effects could quite easily change the rank order of providers, thereby causing a shuffling of the star ratings, especially at the middle of the distribution. Such perturbations could arise because the estimated provider effects have a random error; or because of changes in the data or the model.

<sup>5</sup> The 11 compound variables allowed for effects such as the probability of a successful outcome: varies non-linearly with duration of unemployment; impact of unemployment duration varies by age and gender, or with the local unemployment rate; outcomes for disadvantaged groups vary with the number in the SLA receiving assistance.

<sup>6</sup> This involved adding the square of the unemployment rate to the model.

<sup>7</sup> The distribution of star ratings has been altered to have the same mean as that of the provider effects.

**FIGURE 2. DISTRIBUTION OF PROVIDER EFFECTS: RATIO DEFINITION**



⚠️ The estimated provider effects are themselves random variables, since they are derived from a statistical model. This means they are surrounded by a margin of uncertainty, that should be borne in mind in interpreting the results from the model.

***Effect of varying the way the boundaries of local labour markets are defined***

Although this may be a significant issue, assessing it would require detailed adjustments to the data that were impossible in the time available for the analysis.

**4.2 ADDITION OF NEW LABOUR MARKET AND OTHER VARIABLES**

With DEWR’s assistance, we were able to obtain some new variables, that allowed us to make a start in investigating the possible impact on the model. The variables examined were those readily available, rather than necessarily the most important. They do allow us to make a preliminary test of some of the concerns raised during the consultations.

Variables examined were:

?? METRO – a dummy variable = 1 if the SLA in which the participant resides is in the major metropolitan areas of Sydney, Melbourne, Brisbane, Adelaide, or Perth; = 0 otherwise.

*Job Network members have argued that it is more difficult to achieve KPI outcomes in rural and remote regions.*

?? The physical area of the SLA: BIGAREA – a dummy variable = 1 if the area exceeds 2,000 hectares; = 0 otherwise. About 15 percent of SLAs have area above 2,000 hectares.

*Job Network members have argued that both the size of the local labour market and the density of jobs within the local labour market could impact upon the probability of obtaining a outcomes.*

- ?? FT – the proportion of employed persons in full-time employment in the statistical region in which the participant resides.

*Job Network members have argued that in areas with a high proportion of part time employment it is more difficult to achieve a KPI outcome.*

- ?? IND1-IND17 - Industry shares of employment for the statistical region, as of November 2001, at the broad industry level.

*Job Network members have argued that the industry mix affects the proportion of part time and seasonal jobs, as well as the skill composition of jobs available. This can affect the ease of achieving KPI outcomes.*

- ?? Additional time series data on labour force, employment and unemployment, by SLA, from the DEWR Small Area Labour Market publication.

*Job Network members have argued that it is more difficult to achieve KPI outcomes in areas with small labour markets (i.e. only a few employers, little turnover and/or choice of jobs). This suggests trying total employment in the area as a possible indicator.*

*There were suggestions that a high degree of seasonality in a local market could affect Job Network member performance. This is measured as the coefficient of quarter to quarter variation.*

*There were also suggestions that there are substantial geographical variations in work culture, that affect ability to achieve outcomes. The local participation rate may be an indicator of this. However, we were not able in the time available to obtain the necessary population data.*

- ?? The Centrelink office at which the participant is registered.

*Job Network members have argued that different Centrelink offices have different practices that may affect the ability of Job Network members to achieve an outcome, independently of other effects included in the model.*

- ?? Regional dummy variables for the 19 JN Labour Market Regions.

*Job Network members have argued that there are significant regional effects.*

The new variables predominantly relate to labour market conditions rather than job seeker characteristics. Variables based on the Labour Force Survey are subject to considerable sampling error, that needs to be borne in mind in assessing the results. There is also the issue of possible inappropriate definitions of the boundaries of local labour markets. Adjusting these could alter the results for affected variables and therefore for the model as a whole.

Several of the new variables are dummy variables taking the values, 0 or 1, depending on a particular criterion being satisfied or not. Such variables are open to the criticisms<sup>8</sup> that they are only crude indicators of the underlying economic effects that they purport to measure - also there is a risk that they correlate with some influence other than that intended – giving rise to mistaken interpretations. We discuss this issue further below.

### ***Core experiment: addition of variables METRO, BIGAREA and IND1-17***

A core experiment involved addition of the new variables representing: metro/non-metropolitan SLAs; SLAs with large area; and the proportion of employment by industry (both for the statistical region).

<sup>8</sup> We note that several existing variables in the model are dummy variables - and therefore open to similar criticisms.

As a set, the variables were significant in the equations, causing some changes in estimated provider effects - and hence in the ranking and star ratings. The impacts were larger for IA than for JST. For IA about half the providers were affected. For JST, about 40 percent. The vast majority saw only small movements in star ratings.

Adding the variable FT to the model in addition resulted in only a small further improvement in fit.

- ✂✂ This preliminary experiment to assess the impact of including new variables relating to urban/regional differences in labour market characteristics suggests that they do have the potential to alter the estimated provider effects generated by the model. However, in the vast majority of cases the changes in star ratings were small.
- ✂✂ Addition of the variables METRO, BIGAREA and IND1-17 to the model would provide a short-term method of improving the control for urban and regional differences in labour market conditions.

### *Addition of other variables derived from the Labour Force Survey*

We tried adding extra variables based on the Labour Force Survey to the core model as expanded in the previous section. The new variables were the number employed in the SLA and the coefficient of quarter to quarter variation in employment. We were not able to include the local participation rate, for lack of local population data.

The variables included generally did not add significantly to the explanatory power of the model for IA. The experiment was therefore not performed for JST.

We do not regard this experiment as particularly conclusive, since further research is needed into the definition of regional labour market boundaries. Also the variables used are subject to large sampling error. Similar variables based on the 2001 Population Census would be more reliable.

- ✂✂ Further addition of new variables drawn from the Small Area Labour Force publication did not add to the explanatory power of the model for IA. We do not regard this experiment as particularly conclusive, since further research is needed into the definition of regional labour market boundaries and other issues.

### *Addition of dummy variables for JN Labour Market Regions*

As a separate experiment we added dummy variables for each JN Labour Market Region to the models for IA and JST<sup>9</sup>.

Job Network members drew our attention to the substantial differences in average star ratings for providers in different JN Labour Market Regions. Comparing the upper and lower extremes, average star ratings in Melbourne, Adelaide and Western Victoria are roughly twice those in Riverina, Greater Western Australia and the Northern Territory. The issue for the model is: “To what extent do these differences reflect omitted differences in labour market conditions and job seeker characteristics, and to what extent are they the result of differences in provider efficiency?”

Note that algebraically, the mean of all the provider effects in a region closely reflects the average difference in performance of that region’s providers from those in other regions. Hence the provider effects together are capturing all regional differences that are not specifically reflected in other variables in the model. In other words, they will reflect the average net impact of all omitted regional variables, including any representing labour market or job seeker characteristics.

Dummy variables for JN Labour Market Regions would also capture these effects. But they would overcompensate, in that they would also remove from the model any systematic differences in average provider performance across regions. Subject to this caution, we did investigate the impact on the model of introducing regional dummy variables, since it provided an upper bound to possible impacts of omitted regional effects.

<sup>9</sup> The “dummy variable” for a JN labour market region (e.g. the A.C.T.) takes the value “1” for job seekers that are clients of a JN member operating in that region. It takes the value “0” for all other job seekers. There is a different dummy variable for each JN labour market region. Adding them to the model controls for observed differences between regions in the mean probability of successful labour market outcomes.

As might be expected the regional dummy variables have a large impact on the overall fit of the models for IA and JST and on the star ratings derived from it.

- ⚡ Adding dummy variables representing JN Labour Market Regions to the models for IA and JST has a large effect on the fit of the models and on the star ratings derived from them. This is to be expected, given the observed differences in average star ratings (and hence in estimated provider effects) across regions.
- ⚡ The dummy variables are picking up these systematic variations in average provider effects across regions. It is unclear how much of this in fact reflects unobserved differences in regional labour market conditions and job seeker characteristics, and how much genuine differences in provider effects. Hence we do not advocate the use of such dummy variables in the model without further analysis of the underlying reasons for regional differences.

### ***Addition of dummy variables for Centrelink offices***

In similar vein, we also experimented with the addition of dummy variables for each Centrelink office to one of the equations for IA (that for PLACED).

The justification for adding such dummy variables is that JN members claimed that their performance was affected by differing practices followed by individual Centrelink offices. However, Centrelink office areas also form a patchwork covering Australia. Such dummy variables will tend also to pick up any effects that correlate with area – and at a finer degree of detail than the JN Labour Market Regional dummy variables.

It is not surprising therefore that addition of Centrelink dummy variables had a substantial impact on the fit of the equation and on the estimated provider effects and star ratings.

- ⚡ Including variables to represent Centrelink office effects has major implications for star ratings. As with the dummy variables for labour market regions, the dummy variables for Centrelink offices may be picking up a variety of effects (including differences in provider performance) besides those specifically relating to differences in Centrelink administrative practices etc. Hence there is a need for further analysis of the reasons why the impacts are so substantial before any move to include these variables in the model.

### ***Discussion***

To understand further the relationship of the regional dummy variables to the other variables in the model we reviewed the overall improvements in the fit of the equations that resulted from the various experiments. This reveals the following pattern.

- ?? the models examined for IA or for JST explain only a fairly small proportion of the variation in observed outcomes. This is to be expected given the nature of the data;
- ?? adding provider effects to the current models adds considerably to their explanatory power;
- ?? dummy variables for JN Labour Market Regions and the new variables (for Metropolitan /Non-Metropolitan; SLAs with large area; and the proportions of employment in different industries) add very little to the overall goodness of fit when provider effects are also present;
- ?? even with these additional variables present, the provider effects still add considerably to the models' explanatory power; and
- ?? these findings apply both to IA and to JST.

From this we draw the following conclusions:

- ⚡ there is evidence that adding new variables to represent differences in regional labour markets would improve the statistical properties of the model. It is also likely to have some impact on the estimated provider effects and star ratings, tending to reduce regional differences;
- ⚡ it appears that, even if such variables are added, there are still likely to be large estimated provider effects, unexplained by such variables; and

☞ it would be possible to make quick adjustments to the specification of the model that would improve the extent of control for differences in urban/regional labour market conditions. This would involve the addition of the variables: METRO, BIGAREA and IND1-17 to the current models.

### 4.3 SECONDARY EDUCATIONAL OUTCOMES IN THE KPIS FOR INTENSIVE ASSISTANCE

The terms of reference required the review team to investigate the impact on relative performance outcomes of altering the weight given to secondary educational outcomes in the KPIS for Intensive Assistance. DEWR requested that we do this by reducing the weight on secondary education outcomes to zero.

We began by investigating the changes in the basic model and the associated star ratings when secondary education outcomes are excluded from the interim and final outcome performance measures. Then, we considered whether there are any systematic changes in the models, focusing on the cases of location and specialist providers.

We found that excluding secondary outcomes did affect significantly the coefficients of the model. For example, it emerged that males had a higher probability of a primary outcome than they did of a secondary outcome.

There was also a significant impact on provider relativities, and hence on the star ratings. The average movement in providers' ranks was 14. Forty five percent of providers had changes in ratings, predominantly by 0.5 stars.

We also examined the nature of the impacts on metropolitan/non-metropolitan providers, and on specialist providers (to the extent that we could identify them in the data).

Similar proportions of providers experienced changes in star ratings in both areas. But twice as many metropolitan experiencing improvements as deteriorations. This was offset by a small number of non-metropolitan providers that experienced falls of 1.0 star or more.

Given the available data, it is not possible to calculate the change in star ratings for specialists versus non-specialists. This is because, at the level of the Labour Market Region, a JN member can provide both specialist and non-specialist services. Using a crude indicator of specialisation, based on the proportion of a provider's clients that came from each disadvantaged group, there was no evidence that "specialists" had different experience to that of other providers.

☞ Reducing the weight on secondary educational outcomes in the KPIS for IA would affect the star ratings of 45% of JN members – mostly by 0.5 star. There is some evidence that providers in metropolitan areas would receive higher star ratings, while a very few non-metropolitan providers would lose 1 or 1.5 star. Based on the data available, it did not appear that specialist providers would fare worse than other providers.

### 4.4 ASSESSMENT OF NEIS

The analysis of the models for New Enterprise Incentive Scheme (NEIS) differs from that of Intensive Assistance (IA) and Job Search Training (JST). This is because:

- ?? only a relatively small amount of data available is available and
- ?? there is no established set of models for NEIS for controlling for differences in labour market conditions and job seeker characteristics.

The review focuses on the outcomes: the percentage of participants off income support three months after ceasing on the NEIS allowance; and the proportion of participants that left the scheme early. We apply the same basic model that is used in IA and JST; assesses the significance of the variables in the model; calculate performance scores for the KPI; and assess whether the results are statistically significant.

For the methodology to be applied to NEIS without modification, the estimated model needs to provide a useful control for labour market conditions and job seeker characteristics. In addition, the provider effects should be estimated with enough precision to make the results meaningful.

We found that:

- ✂✂ the estimated models for NEIS outcomes have few significant variables. That is, outcomes are not significantly affected by labour market conditions and job seeker characteristics, as currently specified; and
- ✂✂ the distributions of provider effects, as measured by the differences between the actual and estimated proportions, can be estimated using the raw outcomes data, i.e., data with no corrections for labour market conditions and job seeker characteristics. For small numbers of providers with outcomes well above, or below, the average, the provider effects are significantly different from zero.
- ✂✂ There may be sufficient information in the data to warrant assessing small numbers of providers as significantly above or below the average performance. However, application of the full model, as for IA and JST, is not justified, given the small amount of data.

The conclusions of the analysis should be reassessed as more observations become available.

## 5. DISCUSSION AND CONCLUSIONS

### 5.1 BACKGROUND

The model allowing control for labour market conditions and job seeker characteristics is conceptually sound. If accurately specified, it should yield true estimates of the effects of JN members on job seekers' prospects of an outcome, that can be used to generate relative performance rankings and star ratings.

However, consultations with JN members raised several areas of concern with the way the model controls in practice for labour market conditions and job seeker characteristics. These particularly concerned treatment of regional and remote labour market conditions – but also numerous other possible concerns relating to data, job seeker characteristics, and other aspects of labour market conditions and the modelling itself.

Some of these concerns stemmed from lack of understanding of the method and its application. Others relate more to policy or administration than to technical modelling aspects – and hence are outside the scope of the study.

Other concerns reflect more relevant issues that have influenced the technical analysis we have undertaken. That analysis has had two focuses. First, we have reviewed the performance of the model in statistical terms, confining attention to the data set already in use. Secondly we have undertaken some exploratory analysis into the impacts of introducing new variables suggested by the concerns expressed in consultations.

The objective has been to review the adequacy of the current model, rather than to seek to establish new preferred specifications. To achieve the latter might be a substantial exercise in data collection, model formulation and analysis – beyond the resources and time available for the present study.

### 5.2 IMPLICATIONS OF THE TECHNICAL ANALYSIS

With some slight differences in emphasis, our findings apply equally to Intensive Assistance and to Job Search Training. We also anticipate that they would be valid for Job Matching. However, the results for NEIS differ.

The results of the technical analysis are mixed.

- ✂✂ The basic specification of the model is sound, and the estimated relationships explaining the probability of a job seeker obtaining an outcome accord with prior belief. The estimated provider effects are little affected by changes to the way they are specified in the model.

However, we have also found that the results (and the provider rankings) are affected to some extent by changes in the functional form of the model using the existing data set. They are also affected by inclusion of a preliminary set of new variables that allow for urban/regional differences. Most of the resultant changes in star ratings are small (i.e. no more than 0.5 stars).

In addition, we note that the estimated provider effects are themselves random variables, since they are derived from a statistical model. Each is subject to a margin of uncertainty that could in principle be estimated from the model.

☞ This suggests caution in using fine differences in estimated provider rankings and star ratings to drive significant real-world decisions.

We also found that estimates would be substantially altered by the inclusion of dummy variables for JN Labour Market Regions, or for Centrelink offices. However, the economic justification for such variables is not directly apparent. Regional dummy variables, for example, overcompensate since they remove any systematic variation in provider quality across regions, as well as any other omitted regional effects.

The provider effects derived from the model in turn are capable of acting as the equivalent of a set of regional dummy variables. This demonstrated by the finding that adding regional dummy variables to a model, already including provider effects, does not improve the overall fit of the model. Rather the regional dummy variables take over some of the explanation previously ascribed to the provider effect.

Possible changes in the model or data are likely to have sharply differing impacts on the estimated provider relativities, depending on whether they:

- A. affect all providers equally, or randomly; or
- B. vary systematically across regions, or providers.

Examples of the first type of change would likely include: general improvements in data quality; including variable(s) to reflect differing numbers of trading days in each month. Adjustments of this kind are less likely to affect providers to differing extents – and are therefore less likely to affect relativities and star ratings.

Examples of the second type of change include: adding variables that correlate with the differing accessibility of jobs in regional areas; or that measure the characteristics of jobseekers that are geographically concentrated or served by only a subset of providers. Adjustments of this kind will affect some providers more than others – and so alter estimated relativities and star ratings.

We have shown that addition of new variables reflecting concerns raised by JN members may well lead to improvements in the model. However, we have also shown that indiscriminate introduction of new variables may well remove elements of the genuine provider difference.

☞ There is scope to improve the model's representation of urban and regional differences in the short run by including a limited set of new variables, METRO, BIGAREA and IND1-17. We judge that this would allow a useful degree of control for urban and regional differences in labour market conditions, without serious risk of removing a substantial element of the differences in provider performance.

### **5.3 WHERE TO FROM HERE?**

The use of a model such as that developed by DEWR to assess performance is a sound, leading-edge approach to performance measurement. The statistical model used to determine provider relativities has the advantages of objectivity, replicability, and reliance on data generated by the Job Network's operation. These are considerable advantages. There is no obvious replacement approach that is similarly universal, comprehensive and objective.

While conceptually sound and operationally reliable, the model for correcting for job seeker characteristics and labour market conditions is potentially capable of refinement.

- ☞ In particular, given their possible interaction with the estimated provider effects, there needs to be further investigation of omitted potential variables that vary systematically across labour market regions, or with particular categories of JN provider.
- ☞ We suggest that the model could be improved in the short run by adding a limited set of new variables controlling for urban and regional differences in labour market conditions, as described above

There are inherent uncertainties about provider effects derived from a statistical model. These are magnified by the likelihood that some further revisions to the model may prove desirable in due course. This suggests that:

- ⚡ consideration be given to reducing, at least temporarily, the number of star rating intervals by eliminating half-star ratings and using only whole numbers between 1 and 5. This would reduce the tendency of ratings to alter in response to small perturbations in the data or model;
- ⚡ Where possible other information should also be considered, in order to buttress star ratings that reflect rankings that are close to the cut-off values of the star distribution.

More generally we suggest that DEWR review whether there are now other sources of information on relative provider performance that could be used to improve the confidence attaching to the assessments. This is important given the potential use of the star ratings in business decisions.

- ⚡ Accordingly, we suggest a two-track approach:
  - ?? over the medium term, undertake systematic research to improve the data and the model in order to reduce the impacts on estimated provider relativities of possibly omitted influences; and
  - ?? in the shorter term, (A) make limited adjustments to the model to improve the control for urban and regional labour market conditions and (B) consider whether there are any other sources of information that might help improve the confidence attaching to assessments based on the star ratings.

We consider these in turn.

### *Research to improve the data and the model*

The following is a tentative list of suggestions for a longer-term program of additional research and data improvement:

1. Review the definition of the boundaries of local labour markets, to reflect more accurately the area within which a resident may find employment.

A suggestion by a JN member that an objective way to do this would be to examine historical job seeker data to see which employer postcodes corresponded to each postcode of residence seems worth serious attention. Alternatively, boundaries could be reviewed directly (but perhaps less objectively) based on local knowledge of feasible travel patterns.

2. Commission research to create a theoretical model of the probability of successful job search outcomes in labour markets of different sizes, geographical dispersion, variety of employment opportunities and rate of structural change. Use it to inform selection of variables and establishment of prior beliefs about the direction and size of impacts on the probability of successful outcomes.
3. Commission research to create a theoretical model, and an econometric framework, that will allow supply side influences out of the control of JN members – such as variations in the practices of different Centrelink offices – to be included rigorously in the model.
4. The imminent release of data from the 2001 Census of Population offers an excellent opportunity to obtain accurate data relating to small area labour markets. These data should be used to explore hypotheses relating to influences relating to local labour markets that could not be definitively explored using less reliable data based on ABS Labour Force Surveys.
5. Review the suggestions arising in the consultations about other impacts on local labour market conditions, that would require the creation of variables based on data from other sources. Determine which of them could be investigated cost effectively.
6. Review the full set of information on individual jobseekers available through the JSCI and Centrelink. Use the available information more fully, along lines suggested in the consultations. Where inaccuracies in Centrelink data seem likely to limit its usefulness, review the cost effectiveness of devoting more resources to improving the quality of at least the aspects of the data that are most relevant to the performance assessment model.

7. Review the variables representing job seeker characteristics in light of comments made in the consultations. Formulate revised variable specifications that more accurately reflect the probability of achieving successful outcomes.
8. Review the extent to which distortions may be introduced into the results through possible inappropriate treatment of certain job seekers, for example those that move residence or die while in assistance.
9. Conduct preliminary experiments to determine the likely make up of the new data set using comparatively simple econometric techniques. As the data set is refined, progressively increase the econometric sophistication. Seek to repair any gaps in the data that prevent full exploitation of interrelationships between outcome variables at different stages in the assistance process.
10. Review also the way that provider effects are represented in the model, so as to increase the robustness of the estimates. For example, investigate possible cross correlations between provider effects for outcomes at different stages in the assistance process, or reflecting common organisational links across JN regional boundaries.

### ***Other possible sources of information***

We can identify three alternative sources of information that may help improve robustness of assessments of provider relativities in the near term. They are:

?? *prices paid for JN services in different regions.*

These prices are determined in a market that balances supply and demand for JN services. Hence they should provide independent direct information on the expected difficulty of achieving paid outcomes in each area. This could supplement the model based information on the influence of labour market conditions and job seeker characteristics.

Possible difficulties with this information might include: conflict with contractual conditions; the fact that prices for JN services are determined in an imperfect market; and the possibility that use of price information for performance assessment might start to influence future tendering behaviour.

?? *calibration of regional performance differences using information on comparative performance of nationally organised providers across regions.*

Some providers have contracts in more than one region and, as a result, have internal information on the extent of provider effort in each region. Such information, from a panel of suitable providers, and independently verified, might be used to verify (and possibly to adjust) the regional relativities in provider performance generated by the model.

Possible difficulties might include: ensuring that comparisons were truly “like with like”; securing the co-operation of providers; securing independent verification of the data; and demonstrating the validity of any adjustments to other JN members.

?? *reducing the weight given to star ratings based on the model in favour of direct assessments of qualitative and other factors in provider performance, presently excluded from the assessment.*

Contract managers and auditors have access to information on provider performance not directly incorporated in the KPIs. This could be used (A) to buttress the assessments of relativities coming from the model and (B) possibly to introduce into the assessment other aspects of performance, such as quality of service, that may be considered desirable on policy grounds.

## 6. ATTACHMENT A: INDEPENDENT REVIEW OF THE JOB NETWORK PROVIDER STAR RATINGS METHOD: TERMS OF REFERENCE

### 6.1 BACKGROUND

The Job Network (JN) ratings method ranks JN members according to their performance against a range of job seeker participation and outcomes measures. These measures are linked to the Key Performance Indicators (KPIs) and contractual conditions set out in the Employment Services Contract and reflect the extent to which JN members have met their contractual obligation to assist job seekers in obtaining employment or education and training outcomes.

On the basis of their ranking relative to other service providers, each JN member is attributed a “Star” rating. Ratings range from “1 Star” to “5 Stars” with half Star increments. JN members are rated for each of the services they deliver in the regions in which they operate.

As part of JN’s competitive market framework, Star ratings are released publicly for each of the 19 JN regions with the aim of informing the individual job seeker’s choice of service provider and informing the wider public about the relative performance of JN members. Ratings for the 137 Employment Services Areas (ESAs) that make up the JN regions are released to JN members only.

#### *Comparability of JN Member Ratings*

The ratings method was initially developed in 1999 with the assistance of the Universities of Flinders and Adelaide, South Australian Centre for Economic Studies (SACES) for use in the assessment of tenders for the second JN contract period. The method was further refined in 2000 with assistance from SACES to provide for the rating of JN members’ performance over the course of the contract.

The ratings method was designed to allow for direct comparisons between JN members operating in different JN regions and ESAs. When assessing achievement against individual performance measures, regression modeling is applied to control for the potential impacts of differences in:

- ?? the labour market conditions under which JN members may operate: and
- ?? the characteristics of the job seekers they assist.

Ratings are also derived using an additional measure of JN members’ success in assisting job seekers from three specific disadvantaged groups: indigenous people; people with disabilities; and people from non-English speaking backgrounds. This equity related measure, which reflects the particular emphasis the Government places on providing assistance to these disadvantaged job seekers, is not subject to control through regression. Each performance measure is given a relative weight and then a combined ranking from all performance measures based on the weights is used to produce the “Star” rating.

#### *Performance Measures*

The National Employment Services Association (NESA), the representative industry body for Job Network, and the National New Enterprise Incentive Scheme Association were consulted in the development of the ratings method. All individual JN members were also invited to provide feedback on the weightings applied to the different performance measures for each service. The table below details the measures used in deriving performance ratings for the Job Matching, Job Search Training, Intensive Assistance and New Enterprise Incentive Scheme (NEIS) services.

Service	Performance Measures
Job Matching	<ul style="list-style-type: none"> <li><del>EE</del> Job placements as a proportion of contracted number.</li> <li><del>EE</del> Proportion of placements that are full-time jobs.</li> <li><del>EE</del> Proportion of job placements taken up by: job seekers who have</li> </ul>

	<p><i>been unemployed for more than 6 months; and for more than 12 months.</i></p> <p><i>Proportion of job placements taken up by: people with disabilities, people from non-English speaking backgrounds, and Indigenous peoples.</i></p>
Job Search Training	<p><i>Proportion of job seekers placed in jobs within 3 months of their participation in Job Search Training.</i></p> <p><i>Proportion of job seekers for which a JST bonus payment was paid.</i></p> <p><i>Proportion of job seekers no longer receiving income support payments 3 months after their participation in Job Search Training.</i></p> <p><i>Proportion of job placements for Job Search Training participants taken up by: people with disabilities, people from non-English speaking backgrounds, and Indigenous peoples.</i></p>
Intensive Assistance	<p><i>Proportion of Intensive Assistance participants placed in jobs of a minimum 13 consecutive weeks' duration – and resulting in income support payment reductions of at least 70 percent.</i></p> <p><i>Proportion of Intensive Assistance participants placed in jobs of a minimum 26 consecutive weeks' duration – and resulting in income support payment reductions of at least 70 percent.</i></p> <p><i>Proportion of Intensive Assistance participants receiving job placements.</i></p> <p><i>Proportion of job placements for Intensive Assistance participants taken up by: people with disabilities, people from non-English speaking backgrounds, and Indigenous peoples.</i></p>
New Enterprise Incentive Scheme (NEIS)	<p><i>Number of job seekers commencing in NEIS as a proportion of contracted commencements.</i></p> <p><i>Proportion of NEIS participants no longer receiving income support 3 months after the end of programme participation.</i></p> <p><i>Proportion of NEIS participants withdrawing from NEIS and returning to income support.</i></p>

For all four services, only those job placements of sufficient duration to attract Job Network fees are counted for performance assessment purposes. Job Network members may place eligible job seekers into additional short-term temporary or casual jobs that, due to their brief duration, do not attract fees.

Initial ratings have been produced based on early data for NEIS. The ratings method for NEIS will be refined and finalised once sufficient programme commencements and outcomes data are available.

It is planned to use the rankings of JN members produced through the performance assessment method as part of any future JN purchasing process in line with the competitive market framework. The Department's aim is to ensure that the assessment method represents current best practice and that there is a common understanding within the industry of the assessment method.

## 6.2 AIMS AND OBJECTIVES OF THE REVIEW

With a view to ensuring current best practice, the Review will assess the way in which the method used to rank JN members against the current performance measures for each service accounts for differences in labour market conditions and job seeker characteristics. To inform potential future ratings development for Intensive Assistance in particular, the Review will examine the method used to combine employment and education outcomes to produce composite performance measures. The performance measures for all services are linked to the overall objectives of

---

JN and the related contractual KPIs and therefore fall outside the scope of the Review. The relative weights of each performance measure are policy matters and so would also be outside the scope of the Review.

The specific scope and objective of the Review is to:

- ?? identify any potential for improvements in the control for differences in labour market conditions and job seeker characteristics, including substantive guidance on the implementation of any such refinements; and
- ?? examine the impact on relative performance scores of any change to the counting of Intensive Assistance secondary education outcomes to produce interim and final outcomes performance measures.

As part of the Review, consultations with NESAs and JN members will be conducted to identify areas for potential refinement of the ratings method and issues that have emerged within the industry associated with the use of the ratings system. This will inform the Review process and also identify ways of improving JN members' understanding of the assessment model and the way it is used to monitor performance. Related consultation sessions will be held in all capital cities plus the regional centres of Orange, Newcastle, Townsville, Bendigo and Alice Springs.

In addition to producing reports, the Review team will provide observations to the Steering Committee (which incorporates representatives from the Department of Employment and Workplace Relations and NESAs) on key issues emerging from the consultations related to the use of the Star Ratings system

### **6.3 TIMELINESS AND REPORTING**

The Review will be conducted in three phases.

- ?? Initial technical familiarisation with the regression method and broader Star ratings system.
- ?? The conduct of consultations with JN members.
- ?? Production of draft and final reports to the Steering Committee.

An interim progress report is to be provided on 25 January 2002. The final draft report is to be provided by 21 February with a view to completion by 28 February 2002.