

**Disability Employment Services
Performance Framework**

Industry Reference Group

REPORT

18 August 2009

Contents

1. Executive Summary	2
2. Background	5
3. Introduction	5
4. Benchmarks	6
5. Comparative Performance model	7
5.1 Regression.....	8
5.2 Ratings bandwidths.....	9
6. Key Performance Indicators	10
6.1 KPI 1 - Efficiency.....	10
6.2 KPI 2 – Effectiveness.....	11
6.3 KPI 3 Quality – Disability Service Standards.....	12
6.4 Performance measures and weights.....	12
6.5 Weightings within programs.....	14
7. National Panel of Assessors	14
7.1 KPI 1 – Efficiency.....	14
7.2 KPI 2 – Effectiveness.....	14
7.3 KPI 3 – Quality.....	15
7.4 Performance measures.....	15
8. Future development	16
9. Conclusion	16
Attachment A – Industry Reference Group Members	17
Attachment B – Terms of Reference	18
Attachment C – Impact Analysis	19
1. Impact of Wages, Hours and Maintenance Component of DEN Capped.....	19
2. Impact of Removing the Fixed Distribution on Star Ratings.....	21
3. Illustration of Proposed KPIs and methodology on current providers.....	23
4. Job in Jeopardy further analysis.....	25

1. Executive Summary

This is the Report of the Industry Reference Group (IRG) established by the Minister for Employment Participation, the Hon Mark Arbib MP, to provide advice on a performance management framework for the new Disability Employment Services (DES), commencing on 1 March 2010. This framework is designed to support quality services and results. It responds to clear feedback received during public consultations on the need to improve the current approach to performance management for Disability Employment Network (DEN) and Vocational Rehabilitation Services (VRS).

It was clear to the Group that the performance management system currently in place will need significant revision to align it with the Government's new policy objectives and the intended reforms to service delivery, including –

- uncapping services – the new demand driven model for Programs A and B will mean all eligible job seekers will have access to services and providers will be required to accept all eligible job seekers
- clear focus on employment outcomes, and
- new incentives for training and skills development.

As a result, the IRG proposes a number of changes to the current performance framework. These are intended to update current practices to reflect the design principles and incentives built into the new disability employment services and to improve their transparency and simplicity.

In responding to its terms of reference, the IRG has considered contributions to the review of Disability Employment Services including feedback on Discussion Papers released in September and December 2008 and the Exposure Draft of the New Disability Employment Services and Employer Incentives scheme 2010–2012 released in May 2009. The IRG has also considered information provided by the Expert Reference Group formed to develop a performance framework for Job Services Australia.

Key objectives supporting the IRG's recommendations include:

- a fair system that measures provider effects and produces reliable comparison based on actual performance
- transparency for job seekers and employers as to how providers are rated and assessed
- improving information available for job seekers and employers as they choose a provider
- a consistent focus on achieving outcomes for job seekers supported by clearer and more timely information to assist providers manage performance including identifying what is working and what can be improved, and
- a performance management framework which supports a commitment to continuous improvement.

The IRG gave close consideration to the call from many providers to move towards benchmarking as an alternative system to the current framework. While it was felt that this may be a legitimate approach, particularly given its ability to address many of the criticisms of the current model, some critical elements of a reliable benchmarking system could not be developed ahead of the first contract period.

In summary, the IRG recommends:

Programs A and B

Comparative performance

Provider performance be assessed relative to other providers with consideration to job seeker and labour market circumstances (like with like). Performance be measured by programs and published for providers and job seekers at both the site and ESA level.

Methodology

The system apply a regression methodology to ensure performance ratings take account of the effects of labour market conditions and participant characteristics. Using a well tested regression methodology removes effects that are beyond a providers' control and allows for valid "like for like" comparison of providers.

- It is also recommended that further work is undertaken to ensure the regression method is as simple and transparent as possible without compromising its predictive power.

Distribution of Providers across ratings bands

There should be no fixed distribution of the number of providers reported in each ratings band and no “half stars”. The assignment of providers into each of five ratings bands should be determined by the distance a provider’s performance is from the average performance of all other providers.

- It is recommended that further analysis be undertaken during the first deed period to establish whether in the longer term other approaches such as benchmarking could increase certainty for providers without compromising the integrity of the performance management system or weaken the systems emphasis on continuous improvement.

Measures and Weightings

The achievement of sustainable employment through the measurement of 13 and 26 week full outcomes should be placed at the centre of the performance framework. These measures would be common to both program A and B, however, weightings will take account of differences in the programs, particularly in the provision of Ongoing Support.

An Ongoing Support measure should be included in the framework with more weight given to this in Program B to reflect the importance of this element in that Program.

Referral to Commencement

The KPI 1 (Efficiency) should include a measure of referral to commencement rate. This will support the Government’s decision to uncap the programs and is intended to ensure all eligible job seekers have access to the services they need. This is particularly relevant for volunteers who are not required to participate.

- Further work to be undertaken on the detailed specifications of the framework includes consideration of circumstances in which referrals would not be counted in this measure – for example where the Provider seeks a review of the referral decision and a new referral option is agreed with a Job Capacity Assessor.

Timeliness Measures

A measure of the time taken to achieve a full 13 week employment outcome be included in KPI 1 (Efficiency) to ensure that participants find sustainable work as quickly as possible and to support education and training objectives whilst guarding against the risk of under servicing harder to place job seekers.

The proposed framework removes two measures in KPI 1 (Efficiency) proposed in the Exposure Draft of the Request for Tender: time taken to engage new participants and time from commencement to being placed in a job. The measure of time taken to achieve a full 13 week employment outcome has been retained. While still addressing Government and consumer concerns that harder to place job seekers may be “parked”, this measure is more consistent with the new design principles which include a stronger emphasis on outcomes and new incentives to consider skills development prior to job placement.

- Further work to be undertaken on the detailed specifications of the framework includes consideration of circumstances in which time in service would be discounted such as during participation in accredited training.

Ongoing Support

An ongoing support measure is introduced to the KPI 2 (Effectiveness). This better reflects the new design principles, however, weightings for this measure should reflect the differences between Program A and Program B.

Measures of Service Quality

Compliance with the Disability Service Standards (DSS) should form the cornerstone of assessments of quality. By better utilising the existing quality assurance framework the IRG believes the new performance framework will avoid duplication and red tape. Feedback from job seekers and employers, compliance with the Code of Practice and Service Guarantees will also be taken into account when measuring provider performance against KPI 3.

- Further work will need to be undertaken to ensure the evidence guides adequately measure all the required quality elements.
- While not explicitly included in the IRG Terms of Reference, the IRG recommends that the Department work closely with the sector to consider opportunities to further reduce duplication between meeting the requirement of the DSS and contract compliance and monitoring activities.

Development of Alternative Future Approaches to Assessing Provider Performance

A benchmarking system would not be feasible at this time. The IRG concluded that the extent of change in the new model will mean existing data could not be used to reliably set performance targets. In addition, technical issues such as a how to balance the desire for some level of certainty for providers ahead of a performance period and the necessity to take account of environmental factors, notably labour market conditions and client characteristics, which are not predictable in advance, will need further investigation.

- The IRG recommends that the feasibility of the future use of benchmarking to assist Providers measure performance be examined. Evidence to either support or not support this option would need to be collected prior to the next contract.

National Panel of Assessors

The performance framework will focus on the delivery of timely, quality assessments and be based on three Key Performance Indicators – efficiency, effectiveness and quality. The IRG recommends that the performance of an individual panellist be measured against quantitative targets, wherever possible.

- The IRG recommends setting targets that are both aspirational and achievable and that are considered to be starting points, with the capacity to increase or reduce those targets in the life of the service period, once industry benchmarks are established for the service.

Concluding Remarks

This report identifies a number of recommendations that are proposed for implementation from the commencement of the new Disability Employment Services. It is recommended that further technical analysis and consultation with providers and key stakeholders be undertaken to ensure the robustness and integrity of the performance framework.

To meet the objective of simplifying the performance framework and improving transparency, the IRG recommends consideration be given to ensuring that the details of the final framework are clearly communicated and providers are given an opportunity to clarify issues with the Department, ahead of the new deed start date.

Technical Reference Group

The IRG recommends that a Technical Reference Group including provider representatives be formed to work with the Department to develop and test improvements and simplifications to the regression methodology for the new Disability Employment Services.

2. Background

The Minister for Employment Participation, the Hon Mark Arbib MP established an Industry Reference Group to provide advice on a performance management framework for the new Disability Employment Services contract, to commence from 1 March 2010. The terms of reference and membership of the Group are set out in [Attachments A and B](#).

This report has been made available in time to inform the Request for Tender for the new Disability Employment Services.

The new Disability Employment Services are intended to simplify the currently complex system and ensure tailored assistance to all eligible job seekers. Importantly, the new services will be demand driven and this is a significant change, particularly for existing DEN Capped providers who, unlike DEN Uncapped or VRS providers, have been managing services within a capped program environment for a long period of time.

The payment structure for the new programs has been simplified and includes clear financial incentives to achieve sustainable employment outcomes along with new incentives for training and education and more responsive support for employers.

Feedback on the development of the new model for the new Disability Employment Services has been very positive. Providers, people with disability and their families and employers have welcomed many aspects of the proposed approach including –

- uncapping access to services
- retaining two separate programs
- stronger focus on education training
- clear rewards for achieving sustainable employment outcomes, and
- simpler fee structure and less administration.

Many submissions to the review criticised the current performance management arrangements arguing that the system was too complex and lacked transparency. It was frequently put that an emphasis on speed to placement led to jobs, but not necessarily sustainable job outcomes. The IRG also considered feedback to the review Discussion papers and the Exposure Draft of the Request for Tender to inform its work and deliberations. It is clear that improvements can and should be implemented to respond to a number of concerns expressed in many of the submissions received.

3. Introduction

The IRG has sought to develop a framework that is consistent with the new design principles; is simpler and more transparent than current arrangements; but which also promotes continuous improvements. It has identified the key stakeholders of the performance framework and taken account of the requirements of each group and their individual needs that must be addressed.

- Clients – by making available information that will help inform their choice of provider. This applies to both employers and job seekers.
- Purchaser- the performance management system enables the oversight of purchased services, promotes continuous improvement in performance, fairly assesses performance of individual providers and can support decisions about future purchasing arrangements.
- Providers – enables disability employment service providers to deliver improved performance.

During the extensive Review of Disability Employment Services consultations a number of concerns were raised about the current method for measuring and reporting provider performance. Some feedback has focused on the disincentives created by some Key Performance Indicators, for example that the DEN key performance indicators do not discriminate between part-time and full-time jobs, or in the case of VRS, the speed to placement measure can mean providers feel pressured to place job seekers in the first job available, rather than take time to build skills necessary to achieve a better or best job placement.

While feedback on some of the details of the current performance frameworks for DEN and VRS is mixed, the prevailing view was that the current method of calculating star ratings, in particular the use of fixed distribution, should be replaced. Providers have identified the following negative impacts of the fixed distribution –

- forcing a set proportion of providers into each star rating band, means that the star ratings may not provide an accurate view of comparative performance if differences in actual performance are marginal, this is exacerbated by the adoption of half star ratings
- as they do not reflect actual relative performance, the ratings and information relating to them does not assist providers identify particular measures that would reliably improve performance.

The IRG proposes that the key objectives for the performance framework include:

- a fair system that measures provider effects and produces reliable comparison based on actual performance
- transparency for job seekers, employers and providers as to how performance is assessed and rated
- improved choice and flexibility for job seekers in finding a provider that suits their situation
- accountability of providers to achieve outcomes for job seekers supported by clearer and more timely information to assist providers manage performance including identifying what is working and what can be improved
- support for a commitment to continuous improvement.

The IRG determined that in addition to these objectives the performance framework for the new Disability Employment Services needed to be –

- nationally consistent
- each provider is able to be fairly compared with others (that is comparing like with like), and
- objective.

4. Benchmarks

Feedback received from some provider's calls for the introduction of absolute benchmarks. The IRG noted views that this approach would create certainty for providers as to what level of performance was expected and would be accepted as sufficient to meet contractual obligations. The IRG also noted views that benchmark systems offer transparency and would enhance cooperation and collaboration between providers.

Feedback indicated that the type and form of "benchmarking systems" referred to in feedback varied across responses. For some providers, it was an absolute benchmark system that would express one or more simple, nationally consistent pre-determined performance measures and a required performance level, for example - a 40% expected employment outcome rate or an 85% referral to commencement rate.

Other responses however, proposed benchmarking systems based on either individually negotiated benchmarks which could take account of local labour market conditions, or relative benchmarks that would take account of local labour market conditions with adjustments calculated on a consistent basis across the national network of providers.

To be effective these three potential approaches to performance management would need to ensure that benchmarks were –

- Appropriate – the benchmark should reflect the objectives of the program and take account of the diversity of participants and local labour market conditions.
- Unambiguous – the benchmark must be well defined and the components of benchmarks clearly identifiable – for example a 45% employment outcome rate with at least 10% attracting an education bonus.
- Specific – specified in advance.
- Measurable – widely applicable and independently established to ensure transparency.

The IRG identified a number of critical elements needed to develop and test the usefulness of performance benchmarks to meet these standards. These included reliable historical performance data to be used to set benchmarks and technical solutions to issues such handling the likely need to move benchmarks up or down retrospectively to take account of external factors that influence performance during a performance measurement

period, but that are not known at the start of a performance period.

The IRG concluded that the extent of change to the new programs, most importantly the introduction of a demand driven model, means that reliable performance data against which to set and test benchmarks would not be available for some time into the first contract period. Further the lack of a technical solution to accounting retrospectively for changing environmental factors which impact on performance would weaken the certainty and transparency gains benchmarking might introduce.

Finally, that more time is needed to consider the different requirements of various stakeholders. For example, on the one hand, job seekers, employers and their representatives would look to benchmarks to set high expectations for providers while providers may argue that pitching benchmarks at best practice levels is not practical.

In light of this the IRG concluded that a benchmarking approach would not be feasible in the first contract period. The IRG recommends that the feasibility of the future use of benchmarking to assist providers measure performance be examined.

The IRG, however, considers that significant improvements to the current system will address many of the concerns expressed in feedback to the Review. Essential improvements include -

- the removal of forced distribution. Performance ratings to be based upon a provider's performance and its distance from the mean average performance
- better communication and transparency around the ratings measures
- better targeted reports to support the ratings measures. Reports should be tailored to satisfy provider requirements. The frequency of reports would also be increased to enable providers to undertake continuous monitoring. For example, while ratings would be publicly released every six months, ratings should also be calculated and made available to providers in the middle of each six month performance assessment period.

5. Comparative Performance model

A number of objectives for the performance framework for the new Disability Employment Services point to the importance of a comparative performance model, including -

- assisting clients (both job seekers and employers) choose services that will best meets their needs
- enabling oversight of purchased services, promoting continuous improvement and developing a sound evidence base for future purchasing arrangements
- performance information that enables providers to reasonably assess how they are tracking in comparison to "like" organisations and in comparison to the program as a whole and to support continuous improvement.

Feedback from providers indicates that reliable, like for like performance measures that help them locate their performance within the network of providers contributes to understanding performance achievements and to identifying areas for improvement. The IRG noted information that some providers have created specific networks with "like" providers for this express purpose.

The IRG however, also noted the high level of concern evident in feedback from providers that the current method is complex, lacks transparency, is difficult to understand and as such did not facilitate or enable continuous improvement. In particular, the use of fixed distribution of a set proportion of providers into each of the star ratings has meant that providers find it difficult to interpret performance information or use it effectively to identify areas that would improve overall performance. Providers have also raised concerns that apparently very marginal differences in actual performance have resulted in very different star ratings.

While the IRG recommends that comparative/relative performance data continue to be an important element of the performance framework, it is recommended that the new framework does not include a fixed distribution.

The IRG recommends that the performance of an individual provider be measured and reported relative to the performance of other similar providers in respect of similar clients facing similar labour markets (a "like for like" comparison) and that these be reported at the site and ESA levels. Regression analysis is the best means of enabling fair comparisons of outcome focussed performance.

To ensure comparative performance measures are fair and nationally consistent, the model needs to consider the role

of regression analysis, the relative importance of factors in the analysis and the expression of performance in ratings bandwidths.

5.1 Regression

Feedback from providers and consumers indicated support for a performance framework that was focused on the priorities of the programs and one that clearly measures provider performance against those priorities. To do this fairly, the methodology must identify and account for external factors, outside the control of the provider, that impact either negatively or positively on performance. For Disability Employment Services Program A and B, the key performance measures are based on the number of sustainable employment outcomes achieved for job seekers. Employment outcome rates however are sensitive to a number of environmental factors including local labour market conditions and job seeker characteristics.

The current DEN and VRS programs each use a regression methodology to calculate individual provider performance (provider effect). This statistical technique adjusts raw performance scores to take account of these environmental factors, thereby producing a comparable score.

The following table provides an illustrative example of the degree to which adjustments are made for external factors in the current DEN and VRS programs. The tables are based on the performance period ending 31st December 2008.

The first column shows the star ratings as they are published. The first row gives the performance ranking based on unregressed data using the star ratings fixed distribution.

Purple cells show the number of sites that have the same rank for regressed and unregressed performance; blue cells show the number of sites that have a higher rank from regressed performance compared to unregressed performance; the yellow cells show the number of cells that have a lower rank from regressed performance.

Table 5.1

DEN Capped	Ranking based on unadjusted performance					Total
Ranking based on regressed performance	1-1.5	2-2.5	3-3.5	4-5		
1-1.5	16	9	1	1		27
2-2.5	8	24	23	5		60
3-3.5	4	16	59	34		113
4-5	4	4	30	79		117
Total	32	53	113	119		317

Table 5.2

Uncapped	Ranking based on unadjusted performance					Total
Ranking based on regressed performance	1-1.5	2-2.5	3-3.5	4-5		
1-1.5	14	8	0	0		22
2-2.5	9	19	19	0		47
3-3.5	0	15	44	22		81
4-5	1	4	19	59		83
Total	24	46	82	81		233

Table 5.3

VRS	Ranking based on unadjusted performance					Total
Ranking based on regressed performance	1-1.5	2-2.5	3-3.5	4-5		
1-1.5	21	10	4	0		35
2-2.5	13	26	25	3		67
3-3.5	0	27	55	35		117
4-5	0	4	33	79		116
Total	34	67	117	117		335

Note: due to time constraints and the availability of data, some sites have been excluded for reasons such as merging, splitting and closing of sites and the unadjusted DEN performance data excludes a measure worth up to 5% of the total score.

These tables demonstrate that the results of regression are that between 77-79% of ratings stay the same or increase. The IRG recommends that performance assessment continue to employ a regression methodology. This is essential to ensuring that assessed performance is a measure of provider effects.

In addition and consistent with the current system, a separate regression be estimated for each measure identified for KPI 1 and KPI 2.

While recommending the continuation of the regression methodology, the IRG noted feedback from providers that the complexity of the methodology can make it difficult to reconcile raw performance against the results of the regression and that providers are unable to predict from their raw performance the impact of regression on their rating. The IRG recommends that a Technical Reference Group including provider representatives be formed to work with the Department to develop and test improvements and simplifications to the regression methodology for the new Disability Employment Services.

5.2 Ratings bandwidths

Examination of the criticism of the current performance management arrangements highlights that the fixed distribution method is the basis of much of the dissatisfaction. This method means that a set proportion of providers are placed in each star ratings category, regardless of the degree of difference in actual performance. This in turn makes it difficult for providers to identify the type and extent of improvements that might be needed to improve overall performance ratings.

The IRG recommends that the new performance framework replace the current fixed distribution method. The IRG has considered a number of options and detailed modelling (Attachment C), including detailed analysis developed for the JSA Expert Reference Group and recommends that comparative performance be expressed in bandwidths relative to the average results.

Feedback suggests that the criticism is not just about forcing set proportions of providers into each bandwidth. The use of half star categories has also been criticised. It is argued that attempts to discern subtle differences in performance can only be meaningful if it is assumed that the performance measurement is extremely accurate. Regression is a means of predicting outcomes within a defined range of probability and statistical experts will, however, always concede that while regression represents the best available method for accounting for non-provider effects, no regression model will perfectly predict future performance. The IRG recommends that the new performance framework dispense with half star categories.

The IRG considered a number of options for setting performance bandwidths, including detailed modelling applied to current DEN and VRS performance data. The IRG recommends that the bandwidths adopted for Job Services Australia, be incorporated into the Disability Employment Services Performance framework. This method would see:

- Providers scoring more than 50% below average achieving the lowest performance band / one-star rating.
- Providers scoring between 50% and 20% below average achieving the second performance band / a two-star rating.
- Providers scoring within 20% of the average achieving the third performance band / a three-star rating.
- Providers scoring between 20% and 40% above average achieving the fourth performance band / a four-star rating.
- Providers scoring more than 40% above average achieving the fifth performance band / a five-star rating.

These bandwidths introduce larger categories, thereby taking account of any weaknesses in a regression methodology. For example the middle band includes all providers performing within 20% above or below average. Based on actual performance, this method will ensure that the actual performance of any provider ranked below average will need to be substantially lower than those ranked above them. It was also recommended that the methodology include rounding up within a standard error when a score is close to the boundaries of a bandwidth. It should be noted the proposed method would mean that all providers could be rated as 3 stars.

The following table provides a snapshot of how current DEN and VRS providers would fare using the proposed ratings system using the adjusted performance scores (that is after the current regression model is applied) from the December 2008 star ratings.

Table 5.4 – Current programs

Program	Band 1	Band 2	Band 3	Band 4	Band 5
DEN Capped	9%	19%	44%	13%	15%
DEN Uncapped	12%	18%	39%	14%	17%
VRS	10%	21%	38%	13%	18%

Table 5.5 – Current programs Fixed Distribution

Star Rating	1	1.5	2	2.5	3	3.5	4	4.5	5
	4%	6%	9%	11%	14%	21%	18%	12%	5%

6. Key Performance Indicators

This section considers what performance indicators are required to best capture the policy objectives set by the Australian Government, including the new demand driven model.

The performance framework for the new Disability Employment Services will focus on the delivery of sustainable outcomes. It is proposed that this framework continue to be based on three Key Performance Indicators (KPIs) – efficiency, effectiveness and quality.

6.1 KPI 1 - Efficiency

Efficiency is an important measure for both Purchaser and client. It is in the interests of eligible job seekers that they gain access to assistance and achieve sustainable employment as early as possible. This KPI and its weighting is designed to be assessed alongside those which respond to the importance of sustainability and quality of placements, particularly following education or training.

The IRG has carefully considered this KPI in the light of feedback to the Disability Employment Services Review discussion papers and the more recent Exposure Draft of the Request for Tender. Current KPI 1 measures in DEN and VRS have been criticised as being too “work first” focussed, leading providers to place job seekers in the first job rather than take time to build skills and capacity for a better or best job.

The proposed KPI 1 measures both the proportion of job seekers referred to a Provider who commence and the time taken between commencement and a 13 week outcome. The former recognises the importance of ensuring all eligible job seekers have access to the services they need and the latter moves from speed to placement to recognising the efficient achievement of outcomes.

The IRG recommends removing the two measures proposed in the Exposure Draft of the Request for Tender from KPI 1: time taken from referral to commencement and average time from commencement to job placement. This recommendation is based on the view that these are inconsistent with the policy objectives of the program and do not add a useful measure of provider performance to the overall framework.

The IRG recommends that further technical work be undertaken to identify circumstances in which some referrals may not be counted in this measure – for example when the appropriateness of referrals is reviewed and the referral subsequently be changed and how best to measure “time” should be measured – for example whether the clock should stop for periods in which a job seeker is in education or training.

KPI 1.1: Proportion of referrals who commence in the program.

Objective: Maximises the number of eligible participants who commence in the program, it is particularly relevant to volunteers who are not required to participate.

KPI 1.2: Average time from commencement to the achievement of a 13 week full employment outcome.

Objective: Ensures the participants find sustainable work as quickly as possible at their assessed work capacity. It addresses concerns that some participants may be “parked” or under serviced for a period of time.

6.2 KPI 2 – Effectiveness

The financial incentives for the new Disability Employment Services clearly focus on outcomes; Job Placement, Full and Pathway Outcomes. The IRG recommends that KPI 2 relate coherently to the financial incentives and measure provider performance against the explicit objectives of the program. The IRG however, recommends that these measures be disaggregated to allow for different weightings for 13 and 26 week full, pathway and bonus outcomes, with most weight given to 26 week outcomes as an important indicator of sustainability.

The IRG recognised that the existing DEN KPI 2 measures covering wages and hours reflected important aspirational goal for job seekers and a proxy measure for quality. However, the technical difficulty of applying these as proxy measures of “optimal employment outcome” along with concerns about the quality, reliability and comparability of the data available to measure these were such, that the IRG recommends that they not be included in the new performance framework. In reaching this recommendation, the IRG reviewed analysis on the effect on Star ratings of removing Wages and Hours from the existing data sets (Attachment C) and considered many instances where reduced hours and different wages may be a result of positive choices.

The IRG recommended the addition of an ongoing support measure to recognise the significant investment made in this element of Program B. While it is recommended that the measure apply to both programs, weightings should be adjusted to complement the differences between Programs A and B.

The fee structure in both programs recognises a Job in Jeopardy outcome where a client remains in their job for a period of 26 weeks. Some feedback from current providers has raised concerns that in some cases the best outcome for a Job in Jeopardy client is to either move into a different job or continue to work but at reduced hours. The IRG noted that as Job in Jeopardy was a small element of both programs some providers would not have sufficient data for a weighted measure and recommended this measure be removed from KPI 2.

KPI 2.1: Job Placements

Objective: Maximise the number of participants placed in employment.

KPI 2.2: 13 week Full Outcome

Objective: maximise the number of participants placed in sustainable employment at their assessed work capacity (or education for certain client groups)

KPI 2.3: 13 week pathway outcomes

Objective: maximise the number of participants placed in sustainable employment or education.

KPI 2.4: 13 week bonus outcome and 13 week outcomes for Indigenous Australians

Objective: to provide an incentive for skills development and training relevant to the local labour market and maximising apprenticeships in occupational areas of shortages. Also provides additional incentive to maximise employment for Indigenous Australians.

KPI 2.5: 26 week Full employment outcome

Objective: maximises the number of participants placed in sustainable employment at their assessed work capacity (or education for certain client groups)

KPI 2.6: 26 week pathway outcome

Objective: maximise the number of participants placed in sustainable employment or education.

KPI 2.7: 26 week bonus outcome and 26 week outcomes for Indigenous Australians

Objective: to provide an incentive for skills development and training relevant to the local labour market and maximising apprenticeships in occupational areas of shortages. Also provides additional incentive to maximise employment for Indigenous Australians.

KPI 2.8: Ongoing support

Objective: an incentive to provide ongoing support to a participant to assist them to maintain their employment.

6.3 KPI 3 Quality – Disability Service Standards

In 2002, the *Disabilities Services Amendment (improved Quality Assurance) Act* was introduced as an amendment to the *Disabilities Services Act 1986*. The amendment introduced new quality assurance requirements for all Australian Government funded disability employment services.

Providers of the new Disability Employment Services will be required to be certified as complying with the Disability Services Standards and associated Key Performance Indicators. Providers will need to maintain certification against the Disability Services Standards throughout the contract period to continue their appointment as a provider.

The Standards are central to the quality assurance system. They define the elements of quality service delivery for people with disability. To achieve certification, service providers need to comply with, and demonstrate continuous improvement against, the Disability Services Standards.

There are five main areas that the standards fall into:

- values and principles (having the right approach)
- service outcomes (having the right end result)
- service delivery (having the right organisational processes in place)
- service management
- staff recruitment, employment and training (having the right staff).

By better utilising the existing framework the IRG believes it will guard against introducing duplication and red tape while better supporting continuous improvement in quality services and performance.

Further work will need to be undertaken to ensure the evidence guides adequately measure all the required quality elements.

Feedback from job seekers and employers and the delivery of service in accordance with the Code of Practice and Service Guarantees would also be taken into consideration. Where a provider chooses to invest in additional quality systems such as business excellence frameworks, DEEWR will also take this into account.

It is proposed that from the commencement of the new contract period, comparative performance for the ratings model should be assessed based on the measures outlined above for KPI 1 and KPI 2.

6.4 Performance measures and weights

The following measures and weights are recommended for use in the new Disability Employment Services. The IRG noted that finalising the KPIs and weightings should remain flexible until the work of the proposed Technical Reference Group is complete. It is proposed that the following table, and caveat, be included in the Request for Tender.

The information below is subject to further detailed technical analysis and consultation before it is finalised and may be subject to change.

		Weightings	
Efficiency		Program A	Program B
1.1	Responsive Proportion of referrals who commence in the program	5	5
1.2	Time taken to achieve employment outcome Average time from commencement to the achievement of a 13 week Full Outcome for employment	5	5
Effectiveness			
2.1	Job Placements Proportion of commencements that are placed in employment	5	5
2.2	13 week Full Outcome Proportion of commencements that achieve a 13 week Full Outcome	25	20
2.3	13 week Pathway Outcome Proportion of commencements that achieve a 13 week Pathway Outcome <i>(including remote education-related outcomes)</i>	5	5
2.4	13 week Bonus Outcome Proportion of participants who achieve a 13 week Bonus Outcome <i>(including 13 week outcomes achieved by Indigenous Australians)</i>	5	5
2.5	26 week Full Outcome Proportion of commencements that achieve a 26 week Full Outcome	35	30
2.6	26 week Pathway Outcome Proportion of commencements that achieve a 26 week Pathway Outcome	5	5
2.7	26 week Bonus Outcome Proportion of participants who achieve a 26 week Bonus Outcome <i>(including 26 week outcomes achieved by Indigenous Australians)</i>	5	5
2.8	Ongoing Support Proportion of ongoing support participants who remain in employment or exit ongoing support as an independent worker	5	15
Quality			
3.1	Quality of services Compliance with the Disability Service Standards and DEEWR satisfaction with quality of services delivered	Compliance with the Disability Service Standards and DEEWR satisfaction with quality of services delivered	
3.2	Stakeholder satisfaction Participant and employer satisfaction with quality of services delivered	Participant and employer satisfaction with quality of services delivered	

6.5 Weightings within programs

In considering how to weight the various measures in the model, the IRG has sought to develop a system that is consistent with the design principles (and payment structure) of the proposed new Disability Employment Services contract.

The measures are consistent across both Programs. The weightings for ongoing support are different for Program A and B and this reflects the fact that Program A participants can only access flexible ongoing support. The remaining 10 per cent for program A participants would be allocated as 5 per cent each to full 13 week and 26 week outcomes, this is consistent with the guiding principle outlined in the above paragraph.

7. National Panel of Assessors

The performance framework for the National Panel of Assessors will focus on the delivery of timely, quality assessments for the Supported Wage System, Ongoing Support and Workplace Modifications. It is proposed that this framework build on the existing framework for Supported Wage Scheme and Workplace Modifications Scheme and be based on three Key Performance Indicators (KPIs) – efficiency, effectiveness and quality. The same KPIs and measures apply for all three types of assessment services.

IRG recommended that the performance of an individual panellist be measured against quantitative targets, wherever possible. It is not recommended that any type of relative performance assessment (such as developing a star ratings model) be undertaken for the National Panel of Assessors. The targets for measures in both KPIs 1 and 2 are to be considered as starting points, with the capacity to increase or reduce those targets in the life of the service period, once industry benchmarks are established for the service.

7.1 KPI 1 – Efficiency

KPI 1 measures the timeliness of organisations in undertaking assessments, completing and submitting assessment reports, and correcting assessment reports that require further work. KPI 1 also measures the rate of assessment allocations that are accepted by organisations and, where assessment allocations are rejected, whether or not acceptable reasons are provided for those rejections.

The IRG recommended setting targets that are both aspirational and achievable, and proposed 90%.

7.2 KPI 2 – Effectiveness

KPI 2 evaluates whether or not assessors are undertaking accurate, individualised assessments that respond to the circumstances of each individual client. The IRG recommended using a sampling process to evaluate assessments, particularly for the new Ongoing Support Assessments.

The IRG also considered that the sampling process should be supplemented by consideration of instances where an organisation has had a higher than average number of appealed decisions overturned for an organisation. The IRG discussed the types of behaviour that this measure may influence – for example, there is a possibility that this measure may give assessors an incentive to make assessment decisions that are less than optimal but are less likely to be appealed. On balance, the IRG considered that the measure should be included in KPI 2 as it is an important measure of effectiveness, and in recognition of the impact that assessor effectiveness will have on clients, employers and Program A and B providers.

To ameliorate these concerns, it will be critical to ensure that the appeals process for assessments is robust and involve an informal element that allows assessors and stakeholders to discuss and revise a decision where appropriate before proceeding to a formal review. The IRG recommended that the assessment appeals process be developed in conjunction with the sector prior to the commencement of the services.

In addition, KPI 2 measures the thoroughness of assessment reports, measuring whether or not assessment reports are considered by DEEWR to be complete, and not requiring any further work. The IRG recommended setting a target of 90% for this measure.

7.3 KPI 3 – Quality

KPI 3 measures stakeholder satisfaction with the service. This KPI will be measured by (but not limited to) the results of stakeholder satisfaction surveys, feedback and complaints, where these are not related to program settings. The IRG recommended including the capacity to take into account feedback from other assessors that subsequently assess the same client and may have views about the quality of previous assessments.

7.4 Performance measures

The following measures are recommended for use in assessing the performance of organisations on the National Panel of Assessors. It is also recommended that a caveat be inserted with these measures until they have been considered by a Technical Reference Group. The proposed wording would be:

The information below is subject to further detailed technical analysis and consultation before it is finalised and may be subject to change.

KPI	Description	Measure
Efficiency		
1.1	Timeliness	a) 90 per cent of assessments are completed within the timeframes set out in the Deed. b) 90 per cent of allocated assessments are accepted by the Panel Member. c) Where Panel Members reject allocated assessments, the Panel Member provides acceptable reasons for all rejections. d) 90 per cent of assessment reports are lodged within 7 Business Days of the date the assessment is due to be conducted. e) Where the Department has returned Assessment Reports to the Panel Member for corrective action, all subsequent reports are finalised and submitted to the Department within 5 Business Days of receipt of request by the Panel Member for SWS and OSA and for WMS, within 2 Business Days.
Effectiveness		
2.1	Accurate, individualised assessments	a) Assessments are conducted in a manner that responds to the individual with disability's circumstances, measured by: (i) DEEWR sampling of assessment reports; and (ii) taking into consideration where there is a higher than average level of appealed decisions that are overturned.
2.2	Thorough assessment reports	b) 90 per cent of assessment reports sampled by DEEWR are accepted as complete, without requiring further work.
Quality		
3.1	Stakeholder satisfaction	a) DEEWR's satisfaction with the delivery of the Service, as measured by but not limited to results of stakeholder satisfaction surveys, feedback (including from other assessors subsequently assessing the same client) and complaints.

8. Future development

Next Steps

To fully develop and implement the performance management framework, ongoing consultations will need to occur with providers. These consultations will build on the recommendations of this report. It is therefore recommended that a communications strategy be developed to complement this work.

Issues for finalisation at the earliest possible stage include:

- Establishment of a Technical Reference Group – to contribute to the development of a simple but reliable regression methodology
- Additional development work undertaken on the evidence guidelines to support KPI 3.
- Further analysis and development of the weightings and measurement definitions for the KPIs.
- Further analysis and development on the placement of the bandwidths.

For the longer term (anticipated to be resolved before the next contract period), the IRG recommends:

- The feasibility of the future use of benchmarking to assist Providers measure performance be examined.

9. Conclusion

The IRG has considered a wide range of issues to ensure that the performance management framework provides a robust and reliable measurement system for the next contract period.

IRG recommendations are contained throughout this report. The recommendations have been based on the data and information provided to the Group. Further detailed analysis will be required to ensure that the performance framework delivers the objectives set out in this report.

IRG believes that the proposed performance framework improves the current model and provides incentives to providers to achieve sustainable outcomes for job seekers and, combined with timely performance information for providers, will support a commitment to continuous improvement.

PERFORMANCE MANAGEMENT

INDUSTRY REFERENCE GROUP

Members

- ACE National Inc. (ACE)
- Australian Federation of Disability Organisations. (AFDO)
- Australian Rehabilitation Providers Association. (ARPA)
- Jobs Australia Limited.
- National Disability Services. (NDS)
- National Employment Services Association. (NESA)
- Department of Education, Employment and Workplace Relations.

Terms of Reference for the Industry Reference Group

The Expert Reference Group will examine and provide advice on future arrangements for performance management of disability employment services that:

- promotes continuous improvement in service and outcomes
- is simpler and more transparent than current arrangements
- enables valid and reliable comparisons of providers across Australia
- rewards sustainable employment outcomes, particularly for the most disadvantaged job seekers
- rewards individualised engagement and guards against parking job seekers
- emphasises lasting benefits and reduces job seeker churning
- encourages better skills development, education and training that leads to sustainable employment outcomes
- increases the focus on employer needs
- includes job seeker feedback and supports job seeker choice
- supports the social inclusion of job seekers including through partnerships and local solutions

Additional analysis of impact of the proposed changes

While it is not possible to reliably predict performance levels in the new programs, the IRG asked that the proposed model be tested against the existing programs, DEN and VRS, performance data. This analysis was undertaken as a way of identifying any unintended consequences. In particular, the IRG looked at the impact of changing the treatment of maintenance on the current DEN performance ratings, the impact of moving to a relative performance based on distance from average and whether there were any worrying trends in the impact of the overall changes on a particular type of providers, including specialists and generalists; location of provider including metropolitan, regional, rural and remote and on large and small organisations.

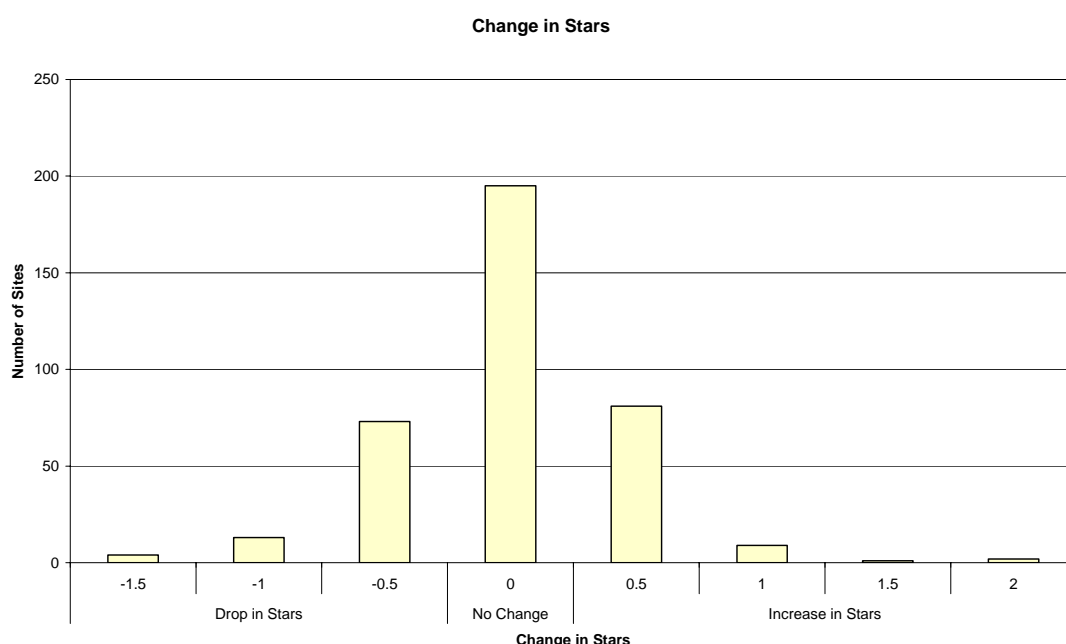
The results of the analysis presented below have generated a level of comfort for the IRG that the proposed new model will not in itself generate any unintended or unwelcome consequences.

Program data for DEN and VRS from the December 2008 performance period has been used. In the case of DEN Capped, performance ratings have been adjusted to account for changes in the Maintenance KPI and the removal of wages and hours.

1. Impact of Wages, Hours and Maintenance Component of DEN Capped

The current DEN Capped star ratings are based on an Employment Assistance component and a Maintenance component which are weighted according to the proportions of job seekers in each component. This proportion ranges from 0-100% of the job seeker caseload. The IRG requested additional analysis around the effect of the weight of the maintenance component, being 15%, as is reflected in the ongoing support KPI. The IRG requested analysis around the impact of Wages and Hours on star ratings so they have also been excluded from this data set. These are based on the performance period ending 31st December 2008.

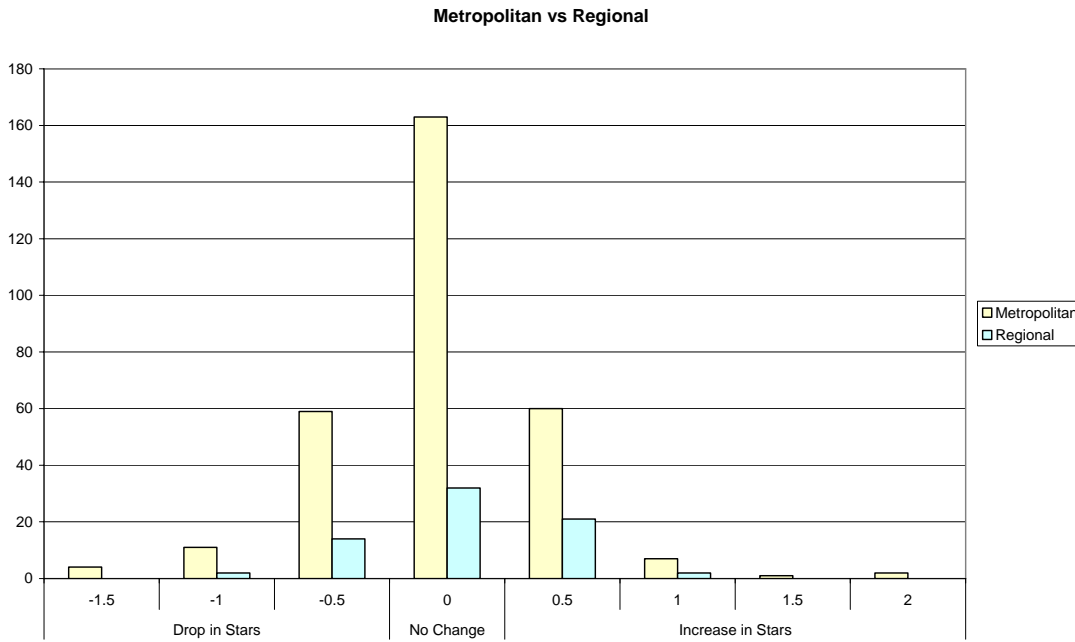
Fig. 1.1 DEN Capped – Impact of wages, hours and maintenance



The chart shows the change in stars from removing wages and hours and fixing the maintenance component. The first point to note is that 92% of sites have a small change (half a star) or no change at all. Data also indicates that the proportion of the Maintenance component is not driving how the site changes.

The following charts show how the above changes are not impacted by such factors as being Metropolitan or Regional, Large or Small, Specialist or Generalist, or by have a high or low DPI caseload.

Fig. 1.2 DEN Capped – Impact of wages, hours and maintenance by site location



This chart shows that the changes are not specific to Metropolitan or Regional sites. In response to a request for further analysis from the IRG, the Metropolitan and Regional categories were disaggregated into the ARIA codes for each site and similar results were obtained.

Fig. 1.3 DEN Capped – Impact of wages, hours and maintenance by site size

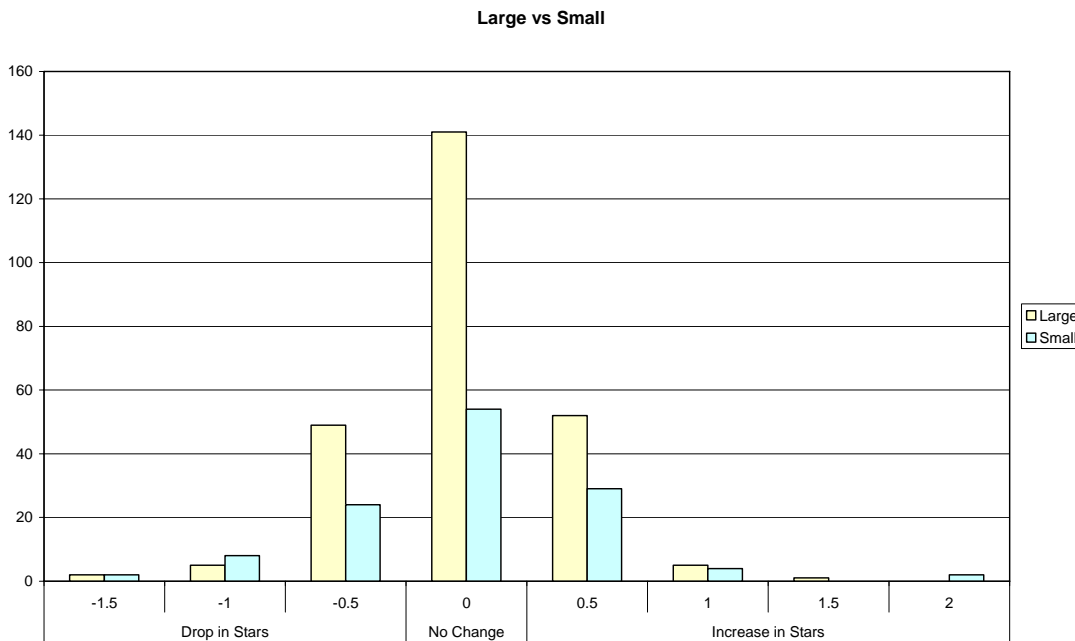


Fig 1.4 DEN Capped – Impact of wages, hours and maintenance by DPI

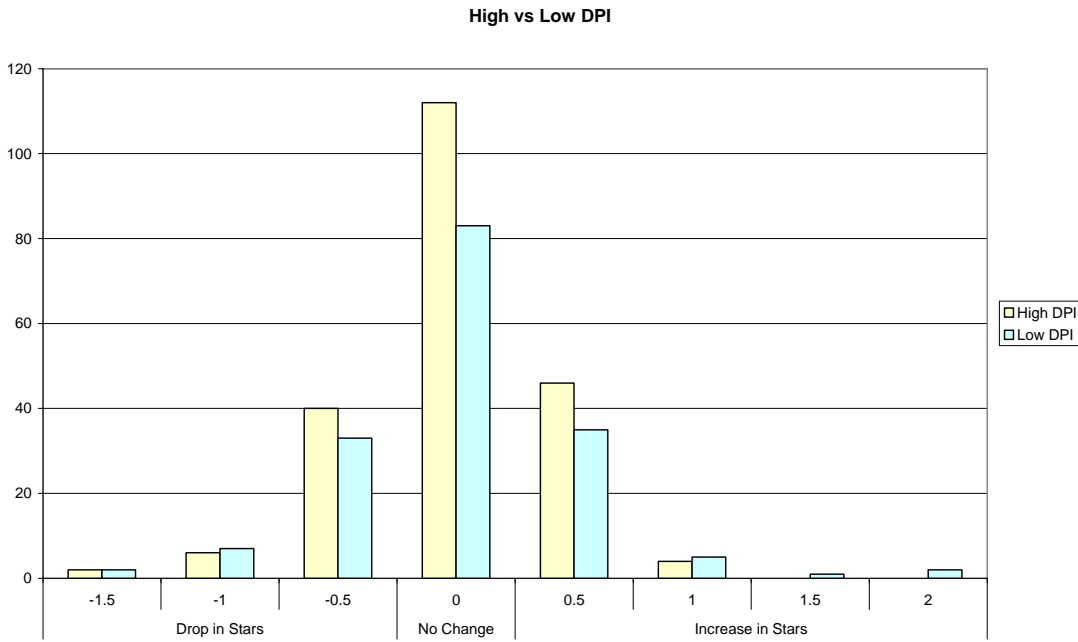
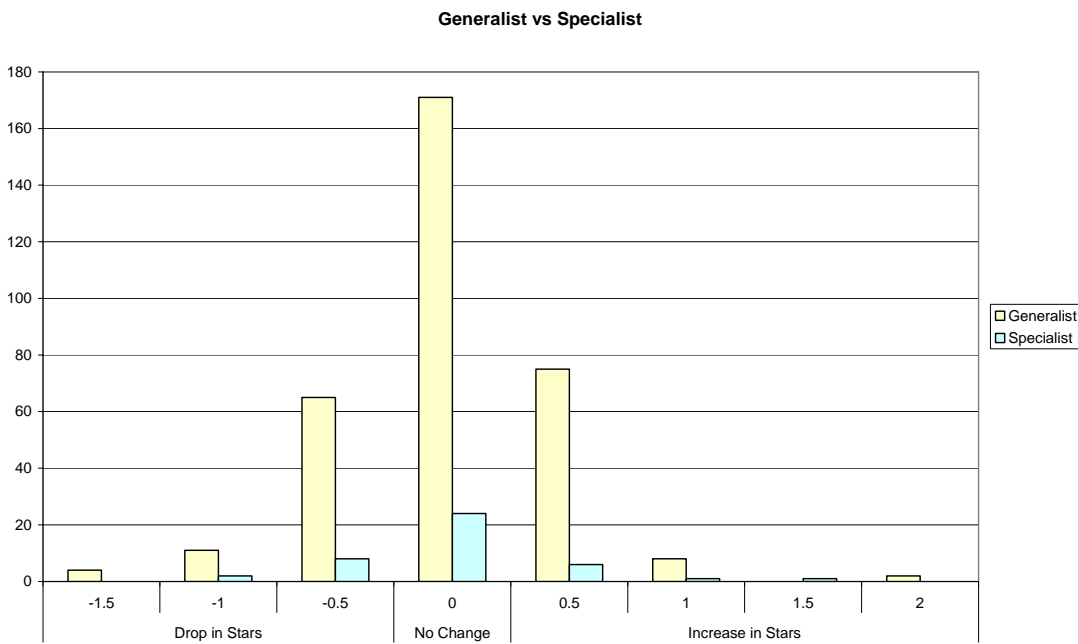


Fig 1.5 DEN Capped – Impact of wages, hours and maintenance by service type



Overall, these charts show that no particular group of DEN Capped sites investigated is advantaged or disadvantaged by removing wages and hours and fixing the weighting of the Maintenance component at 15%.

2. Impact of Removing the Fixed Distribution on Star Ratings

The impact of removing the fixed distribution of sites was investigated for each of the programs. This was done in two stages. First the half star bands were split evenly between the adjacent ratings. For example, in the current methodology, 1.5 stars account for 6% of sites. Under this change, 3% of sites would go to 1 star and 3% to 2 stars. The second stage was to remove the fixed distribution completely and show the numbers of sites in the proposed ratings bands based on distance from the mean. This data is based on the performance period ending 31st December 2008.

Fig. 2.1 DEN Capped – Impact of removing fixed distribution

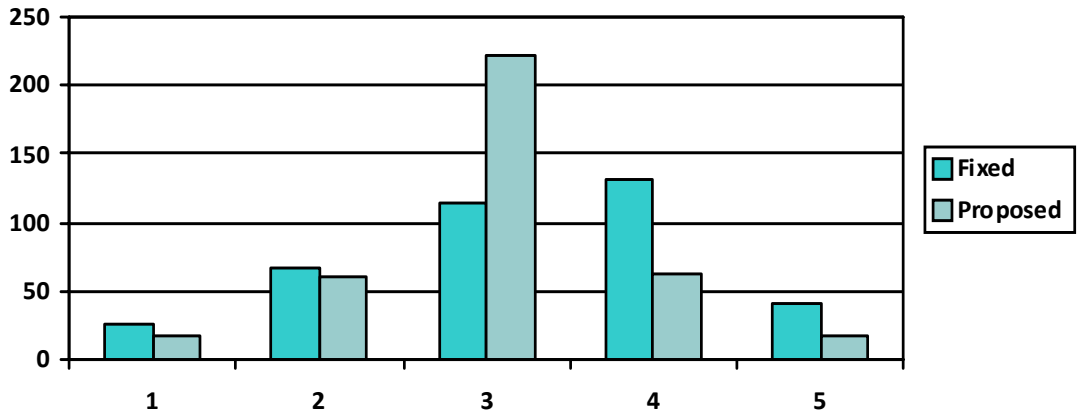


Fig. 2.2 DEN Uncapped - Impact of removing fixed distribution

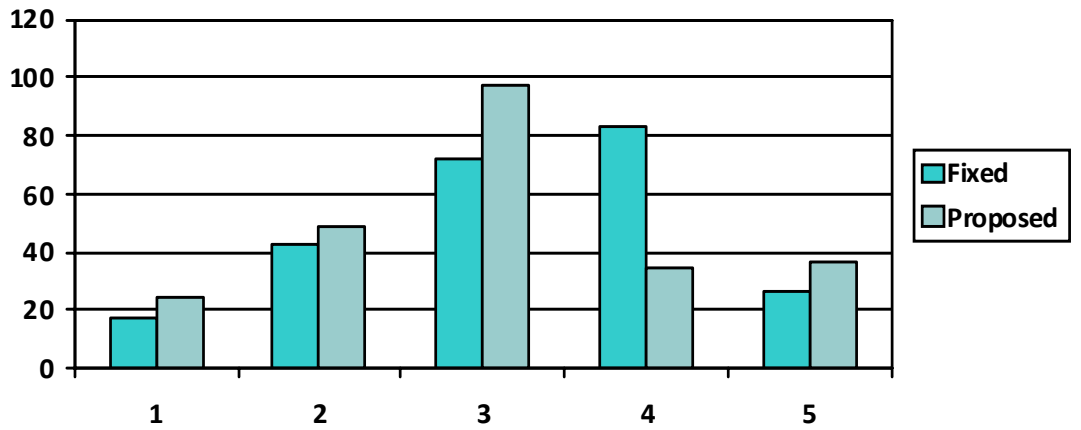
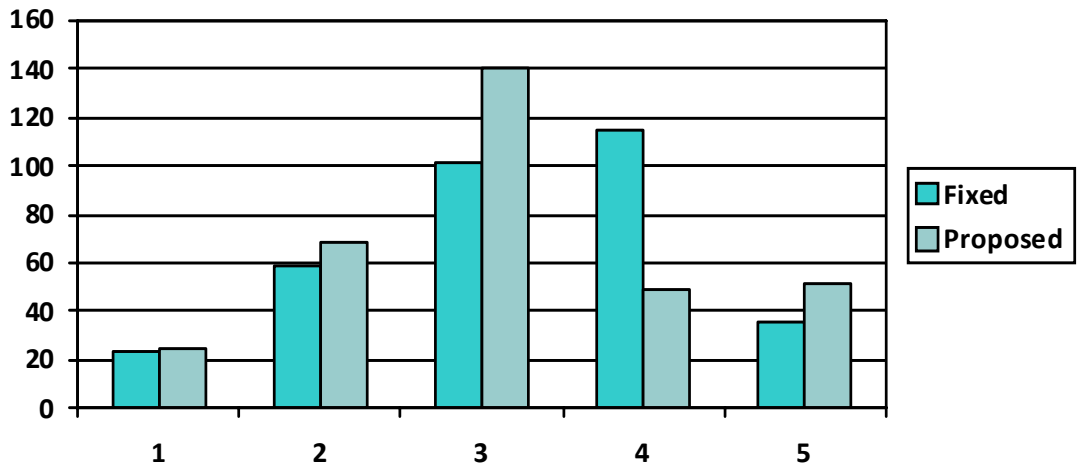


Fig. 2.3 VRS - Impact of removing fixed distribution



The first point to note is that the fixed distribution was positively skewed so that on average, sites had higher ratings. The proposed model does not skew results so on average, ratings are shifted down. This is most clear between the 3 and 4 star ratings. It is also clear that DEN Capped ratings are distributed more evenly and are closer to the mean. DEN Uncapped and VRS, on the other hand, are more spread out which leaves fewer sites in the 4 star band than the 5 star band as it has a comparatively smaller size.

3. Illustration of Proposed KPIs and methodology on current providers

The proposed KPIs, weightings and methodology were illustrated to the IRG based on the performance period ending 31st December 2008. Due to time constraints and availability of data some measures required proxies to complete the dataset.

The following charts show the frequency of sites within each of the new ratings bands. The performance is based on the proposed methodology with the proposed weightings for the proposed KPIs. The frequency is based on ranges of scores of width 0.1.

Fig. 3.1 DEN Capped – ratings frequency with proposed model

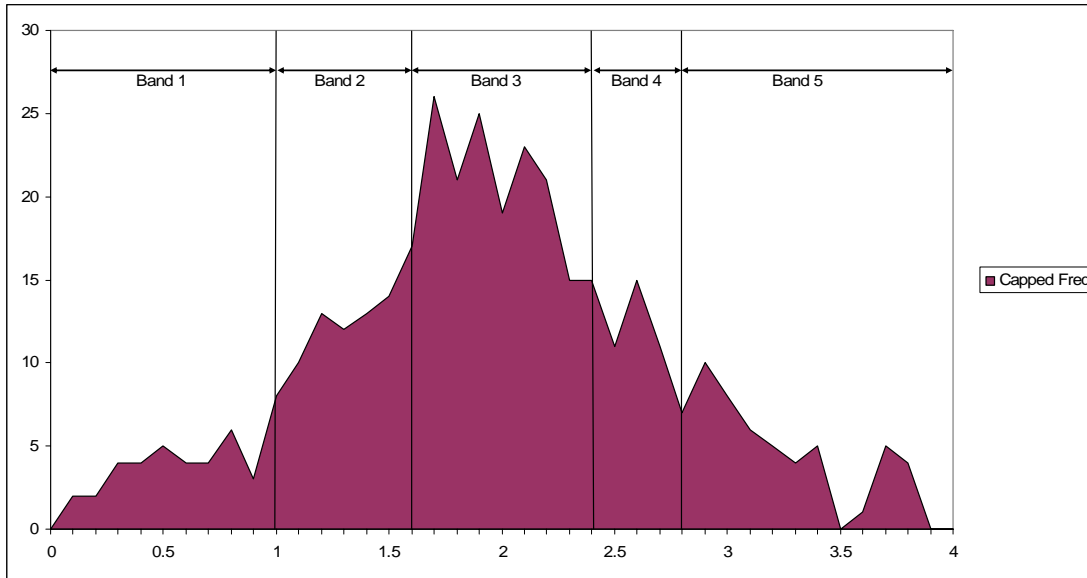


Fig. 3.2 DEN Uncapped – ratings frequency with proposed model

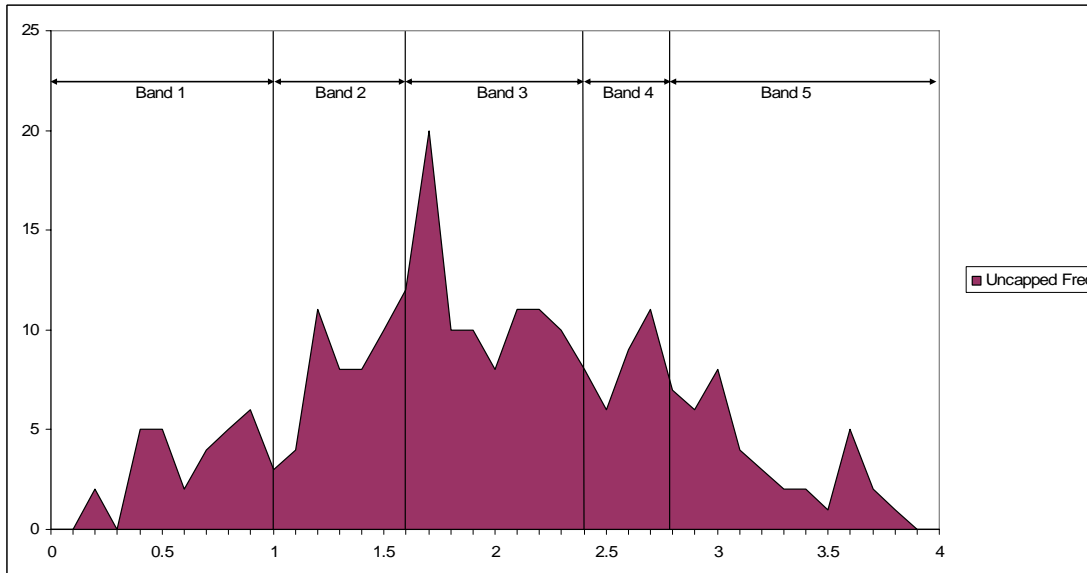
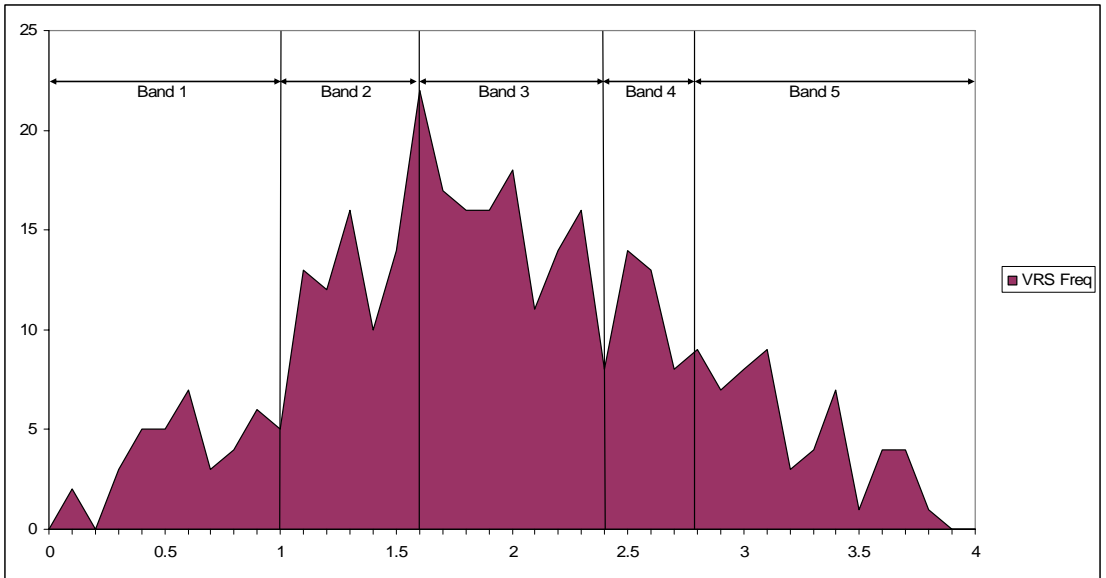


Fig. 3.3 VRS – ratings frequency with proposed model



For each program, the impact of changes between the current KPIs for the program and the new KPIs were generated.

Fig. 3.4 DEN Capped – impact of proposed model

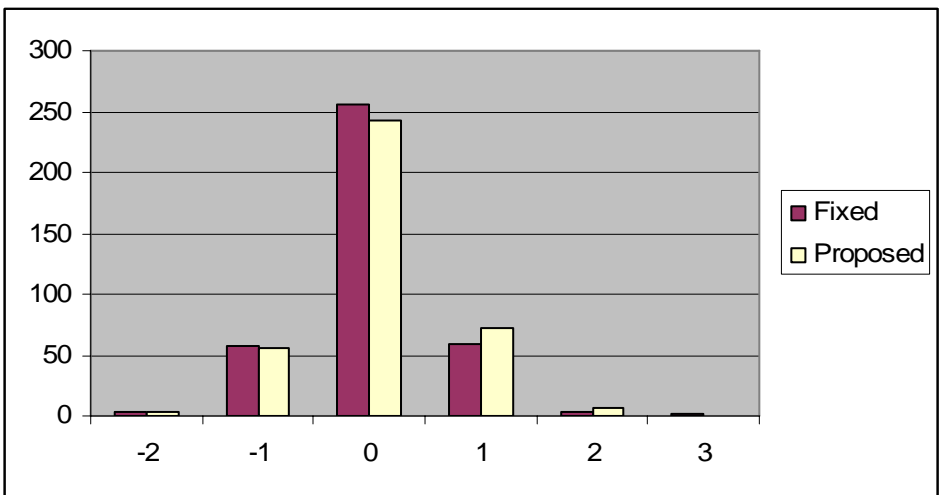


Fig. 3.5 DEN Uncapped – impact of proposed model

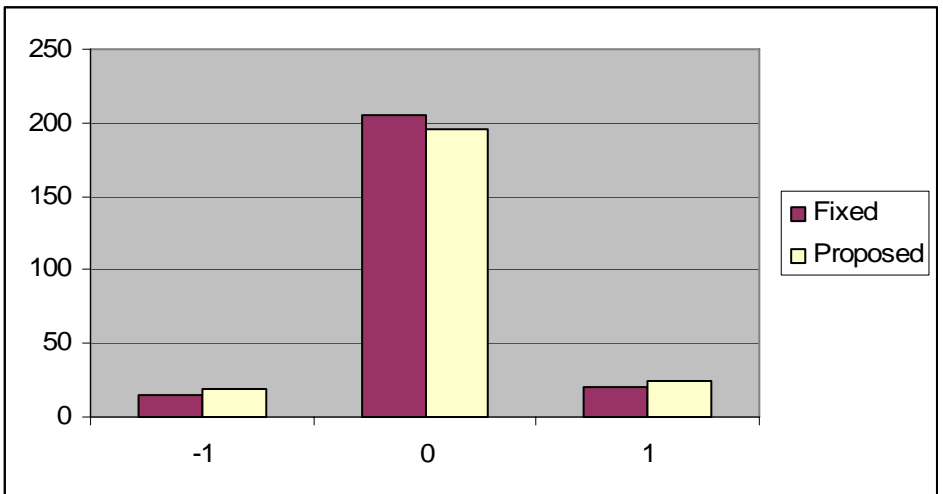
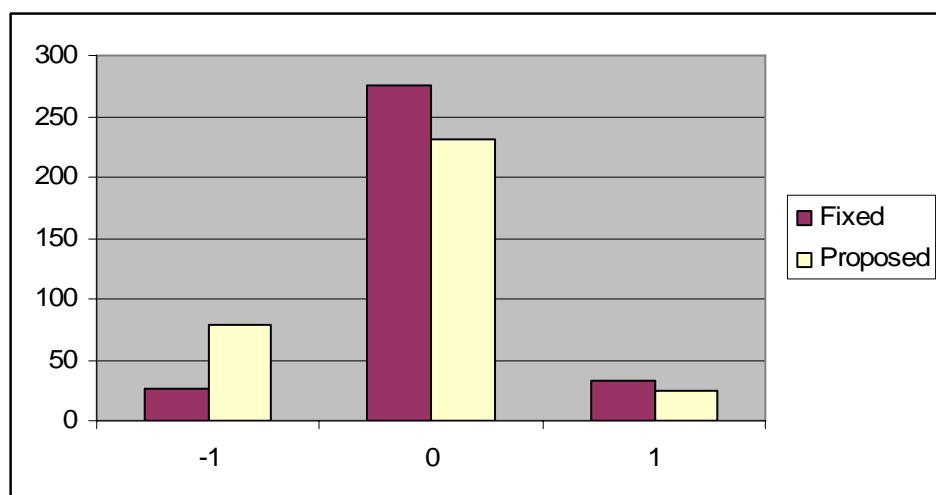


Fig. 3.6 VRS – impact of proposed model



The first point to note is that both DEN Uncapped and VRS only have small changes of up to a star and these make up a small percentage of sites, under both the fixed distribution and the proposed methodology. This is due to the similarity of the KPIs.

DEN Capped overall have some individual outliers in star rating under both methodologies which is consistent with the impact of setting the maintenance component and removing wages and hours KPIs as discussed earlier. However, it is important to note that overall these make up a very small proportion of DEN Capped sites.

The IRG requested that analysis be done to determine whether there was any significant difference in results for different groups including Metropolitan and Regional, Large and Small, High and Low DPI. For VRS, analysis was conducted to assess whether CRS was advantaged or disadvantaged compared with the other providers. The analysis showed that none of the above groups were advantaged or disadvantaged by the proposed KPIs or methodology.

4. Job in Jeopardy further analysis

The IRG recommended the inclusion of Job in Jeopardy outcomes in the key performance indicators be considered. As there are technical difficulties associated with measuring an outcome for such a small client group, it is recommended that a Job in Jeopardy outcome not be incorporated into the performance framework.

Additional analysis undertaken

The main difficulty is that there are too few Job in Jeopardy participants nationally to reliably estimate the probability that a Job in Jeopardy participant will achieve a 26 week Job in Jeopardy outcome. In addition, Job in Jeopardy outcomes cannot be combined with key performance indicator 2.7 (Bonus Outcome) because Job in Jeopardy participants, who already hold employment, are likely to have very different characteristics to non-Job in Jeopardy participants, who are seeking employment. It is considered that in terms of job seekers assisted and the services provided for these jobseekers they are too dissimilar for them to be combined in the one measure (this would lead to different denominators and numerators and would not be feasible). As training is a very important government priority it was not considered feasible to split the weighting across these two measures and attribute the same weighting to Job in Jeopardy.

A further option that was also investigated for including a Job in Jeopardy outcome measure was to combine Job in Jeopardy outcomes with key performance indicator 2.8 (Ongoing Support outcome). This would have resolved the issue of combining two populations with dissimilar characteristics and provide enough participants to reliably estimate the probability that a Job in Jeopardy or Ongoing Support participant will achieve a 26 week outcome. Ongoing Support participants and Job in Jeopardy participants are likely to have similar characteristics, as both cohorts are made up of employed participants who require assistance to maintain their employment.

The major flaw with this approach is that Job in Jeopardy outcomes in Program B (15 percent) would have a higher weighting than in Program A (5 percent), due to the difference in weighting of KPI 2.8 between Program A and

Program B. Due to this difference in weighting it is not considered feasible to introduce Job in Jeopardy outcomes into this KPI.

Supporting Data

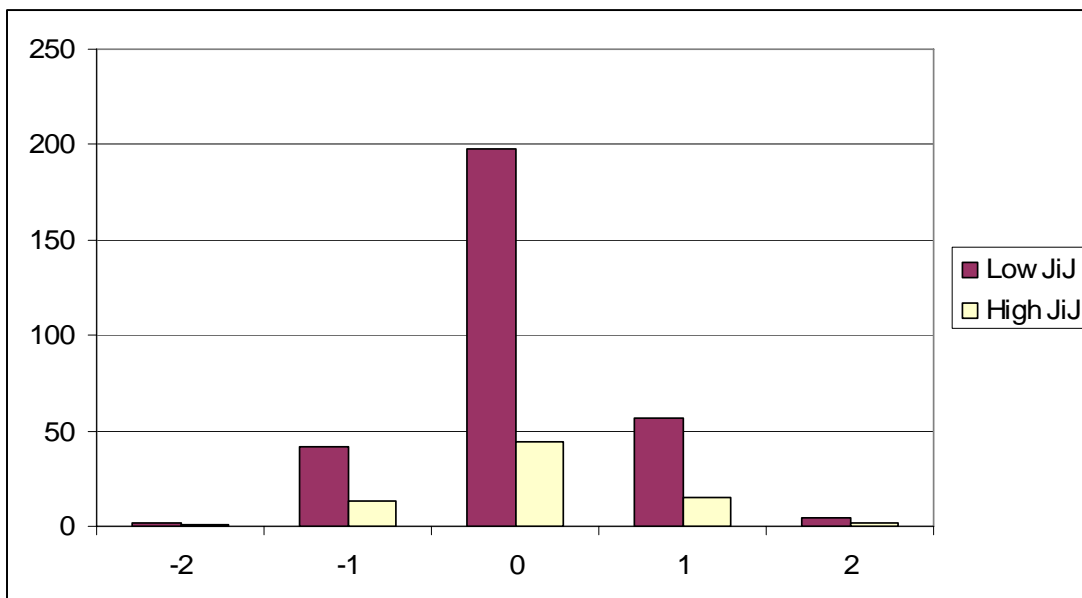
During the period 1st July 2007 to 31st December 2008:

- Preliminary analysis indicated there were 392 Job in Jeopardy participants in VRS and 1,107 participants in DEN Capped and 29 in Uncapped. The DEN data would require further detailed analysis as the way Job in Jeopardy is counted in DEN is quite complex. Job in Jeopardy is approximately 2.19% of the total DEN caseload.

Impact of Job in Jeopardy

The IRG requested additional analysis around the impact of the new star ratings system regarding sites with high caseloads of job in jeopardy. The following chart shows the number of sites who have a high proportion of Job in Jeopardy in their caseload and how their star rating would change under the proposed methodology. For the purpose of this analysis a site is determined to have a high proportion of Job in Jeopardy if greater than 5% of their caseload is made up of Job in Jeopardy. If you were to use a percentage higher than this, analysis would not be reliable due to the low number of sites able to be analysed.

Fig. 4.1 Impact of exclusion of Job in Jeopardy in the proposed model



The graph shows that sites with high proportions of Job in Jeopardy are not advantaged or disadvantaged by the new KPIs for DEN capped. This graph shows the majority of sites do not change rating. There are some sites which move up one star, and a similar number that move down one star.

Analysis for Uncapped indicated that only 2 Employment Service Areas and in Vocational Rehabilitation Services only 2 sites have a proportion of Job in Jeopardy greater than 5% and therefore they have not been represented on a graph. Of these, one goes up a star, one goes down a star and the other two stay the same.