22 February 2006

The measure of central tendency used to set Directlink's regulatory asset base

Australian Energy Regulator

NERA
Economic Consulting

MMC Marsh & McLennan Companies

Project Team

Tom Hird (Ph.D.)

Robert McMillan

NERA Economic Consulting
Level 16
33 Exhibition Street
Melbourne 3000
Tel: +61 3 9623 5020
Fax: +61 3 9245 5123
www.nera.com

Contents

1. Background

The Directlink Joint Venture (DJV) has applied for conversion to regulated status. This requires that the Australian Energy Regulator (AER) sets an asset base for the DJV. It has been determined that the DJV project would not have passed the regulatory test and, as such, the AER has determined to set the value of the DJV assets on the basis of their economic value (EV).

In its draft decision, the AER adopted the median estimate of the EV from the sample of credible scenarios that had been run in the application of the regulatory test[1]. In so doing, it noted that:

> *The AER considers that the best balance to determine an EV that is representative of the credible scenarios is to use the measure of central tendency. Given the range of estimates and the skewed distribution, using a mean to determine a single value is not appropriate because the mean is more affected by extreme values and is therefore not a good measure of central tendency. The median is less sensitive to extreme ranges and this makes it a better measure than the mean for skewed distributions.*
>
> *In determining the median, the ordered middle value is selected when there is an odd number of scenarios. In this case, there is an even number of scenarios. Therefore, the median is determined to be the mean of the two middle values. That is, the mean market benefits of scenarios 3 and 4 results in a median EV of $150.55m for Directlink.*[2]

Responding to the draft decision, the DJV questioned the statistical validity of using the median in measuring central tendency among the sample of credible scenarios and has posited that the mean is a more valid measure in this situation. Their preference for the mean was supported by claims that:

- the use of the mean minimises the risk of estimation whereas the use of the median does not; and
- sampling for the mean has a lower error rate than sampling for the median.[3]

---

1 The regulatory test is defined as the test promulgated by the AER in accordance with clause 5.6.5A of the National Electricity Rules.

2 AER, DJV Application for Conversion and Revenue Cap Draft Decision, 8 November 2005, pages 129-130.

3 DJV, Submission to AER Draft Decision, 9 December 2005, page 18.

2. Statistical measures and sample characteristics

2.1. DJV arguments

The DJV has presented a case for the AER to adopt the mean economic value of the six credible scenarios instead of the median. In doing so the DJV has argued:

1. the AER's approach in adopting the median of the six credible scenarios as an attempt to estimate the median of the underlying population of scenarios; and

2. the sampling error associated with the mean of a **large number of observations** drawn **randomly** from a population is less than the sampling error associated with the median.

The expression of the first point can be found on page 18 of the DJV submission:

> *"The AER appears to believe that the best estimate of economic value is the point which has equal probability of understating or overstating the economic value."*

The expression of the second can be found on pages 18 to 21 culminating in reproduction of two statistical theorems from Freund and Walpole's 1987 text *Mathematical Statistics.* The DJV then relies on these theorems to conclude:

> *"The above theorems imply that sampling distributions of the mean and median both tend towards normal as the sample size increases but the error associated with sampling the median is greater than the error in sampling the mean. The precise difference between sampling errors of the median and mean depends on the underlying distribution—for the gamma distribution, which can be used to approximate the distribution of market benefits, the sampling error of the mean is less than 60 per cent of the sampling error of the median."*[4]

2.2. Summary of NERA's Assessment

If the DJV has correctly interpreted the AER's approach (in point 1 above) as wishing to estimate the population median rather than the population mean then it is difficult to see the relevance of point 2 above. If the AER does wish to estimate the population median then the best estimate of the population median is the median of the sample (assuming that the sample is randomly drawn from the population). The fact that the sample mean has a lower error in estimating the population mean **does not** imply that it has a lower error in estimating the population median. If the AER wishes to measure the population *median* then the statistic it should choose is the statistic that has the lowest sampling error in measuring the *median*.

However, it is possible that the DJV believes that the AER *should* be attempting to measure the population mean and that the sample mean is the best estimate of this (although such an argument is not clearly stated in its submission).

---

[4] DJV, Submission to AER Draft Decision, 9 December 2005, page 20.

It is therefore instructive to ask two questions:

1. "If the AER did wish to estimate the population mean does it axiomatically follow from the statistical theorems presented in the DJV's submission that the mean of the six credible scenarios is the most accurate estimate of the population mean"?

2. "Should the AER wish to estimate the population mean, the population median or some other value?"

The answer to question one is 'yes' provided that the six credible scenarios were randomly drawn from an underlying population of possible scenarios. However, on the basis of the evidence before us this proviso does not hold. When account is taken of how the six credible scenarios were chosen we reach the conclusion that the median of these scenarios is likely to be a better estimate of the population mean than the mean of these six scenarios (see section 2.3).

The answer to question two is a matter for judgement and will depend on the ultimate objective one believes the AER should pursue. If, however, one applies the same standards of certainty to the estimation of economic value as is implied in the legal construction of the regulatory test then it is appropriate for the AER to adopt the sample median (or an even lower value) as the measure of central tendency in the current context. see section 2.4.

2.3. Credible Scenarios Not a Random Sample

The AER's regulatory test identified six credible scenarios. These six scenarios were formed applying two different assumptions concerning the value of unserved energy (USE) and three assumptions concerning demand growth (low, medium, and high).

It can not, therefore, be presumed that these scenarios represent a random sample from the total population of scenarios. This is a point that the DJV acknowledges in their submission. However, the DJV goes on to make an assertion that:

> *The AER has concluded that six scenarios are 'credible'; presumably meaning that the six scenarios should be more representative of the underlying population than had they been chosen purely at random. Credible scenarios should therefore enable better estimates of the underlying population and have lower errors than a purely random sample.*[5]

The statement that the credible scenarios will be *better estimates of the underlying population and have lower errors than a purely random sample* has no clear statistical meaning and the DJV submission does not give it one. The DJV wishes to give each scenario an equal weight in measuring the population mean so it can have no reason for believing one observation is more representative than another. For practical purposes we can only assume that the DJV means each of the six credible scenarios has an equal probability of occurring as would be the case if they were randomly drawn from the population of possible scenarios.

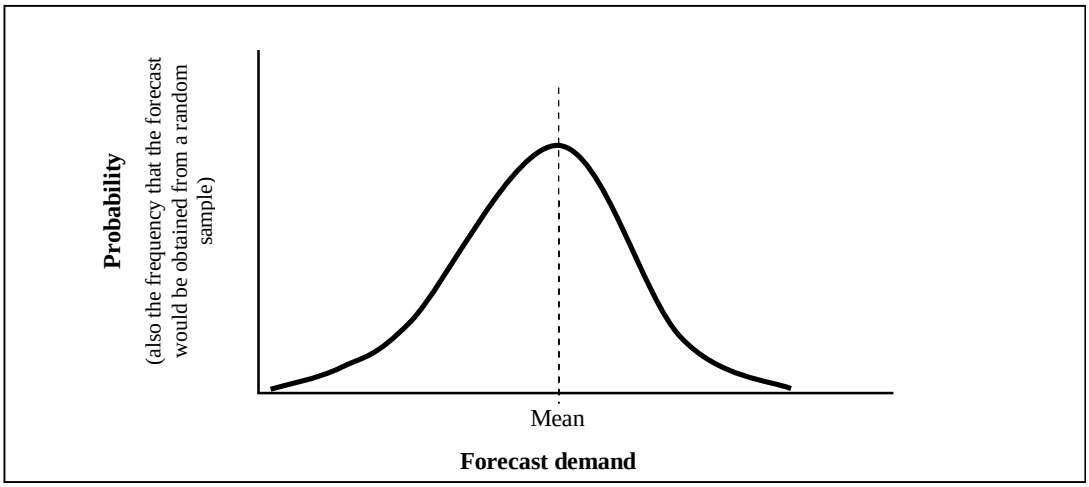
---

[5] DJV, Submission to AER Draft Decision, 9 December 2005, page 19.

We believe that such an assumption is not supported by the facts. The primary driver of differences between the credible scenarios is the assumed forecast of demand. While the demand forecasts used by Burns Roe Worley (BRW) have been verified by Parsons Brinckerhoff Associates (PBA) as reasonable, neither the PBA report nor the BRW report states that the high, medium and low forecasts are of equal probability of occurrence.

In fact, it appears to us more reasonable to assume that the low, medium and high demand forecasts were chosen to reflect the range of credible forecasts - with the medium forecast being the most likely and the high and low forecasts representing the less likely extremes.[6] Figure 2.1 illustrates how a normally distributed population of demand forecasts would look.

Figure 2.1
Normal distribution



From Figure 2.1 one can see that a random sample taken from a normally distributed population of demand forecasts would be expected to return more results that are close to the mean forecast.

The nature of the scenario analysis adopted by BRW (and therefore DJV) deliberately selects demand forecasts that fall in the high, medium and low parts of the demand forecast spectrum. This is noted by the AER who state that:

> *Incorporating each of the low, medium and high growth rates as credible scenarios allows the regulatory test to incorporate the spectrum of likely growth.'*[7]

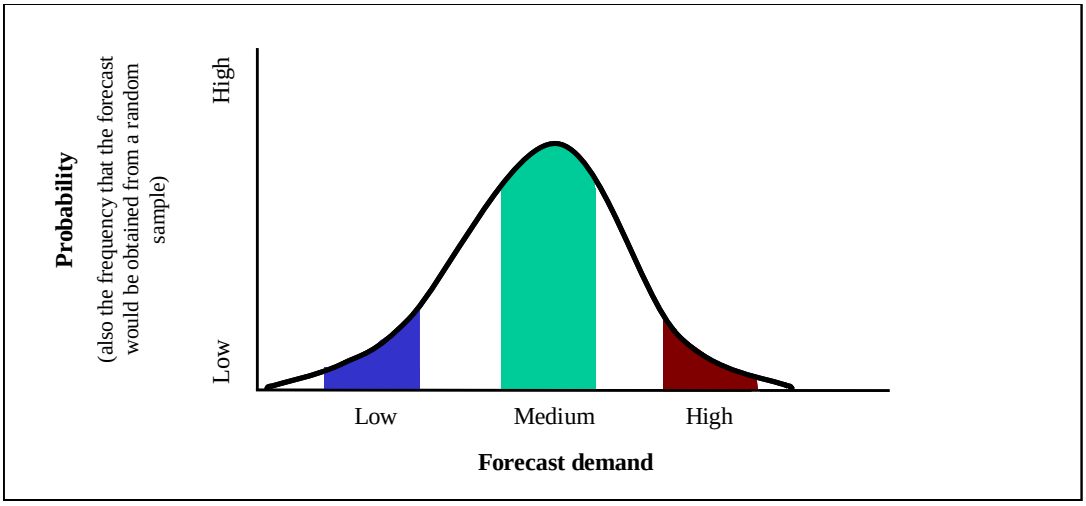
Thus a sample chosen to capture the spectrum of both high and low credible demand forecasts in such a small sample is likely to mean that the observations within the sample are not of equal probability. This is shown in Figure 2.2.

---

[6] We note that these forecasts are sourced ultimately from NEMMCO's Statement of Opportunity which similarly does not state that the forecasts have equal probability.

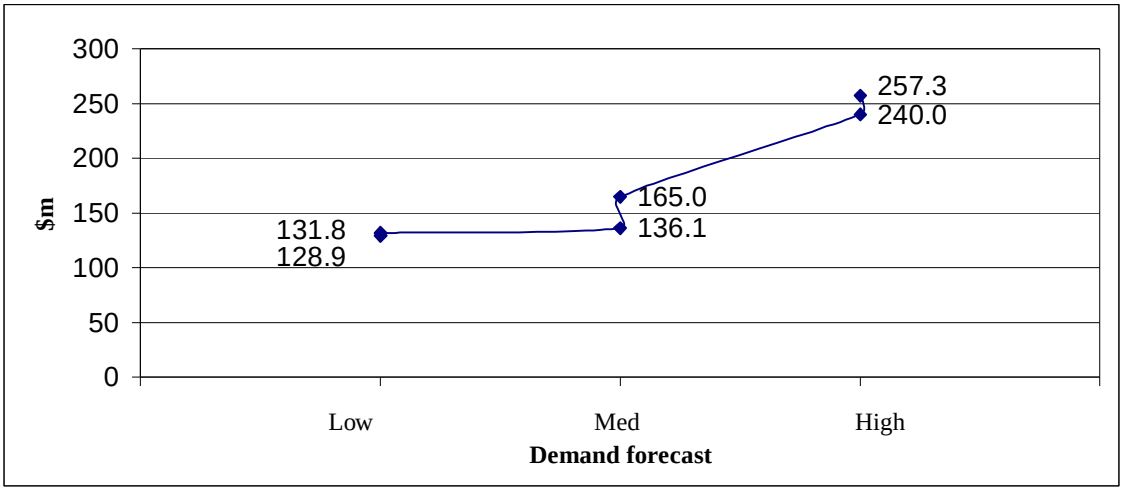
[7] AER, DJV Application for Conversion and Revenue Cap Draft Decision, 8 November 2005, page 118.

Figure 2.2
Spectrum sampling



In other words, the medium credible forecast reflects the most likely credible forecast and the low/high credible forecasts reflect the more extreme (less likely) ends of the spectrum. It would be an error to assume, as DJV does, that they have the same probability of occurring. Such an assumption will be particularly problematic when there is a skewed relationship between estimated economic value and forecast demand. This is indeed the case in the current context as is demonstrated in Figure 2.3 below.

Figure 2.3
Distribution of market benefit scenarios[8]



[8] Sourced from AER, DJV Application for Conversion and Revenue Cap Draft Decision, 8 November 2005, table 11.2, page 129

In such circumstances affording each of the credible scenarios an equal probability will cause an overestimate of the true population mean. This is because the high economic values are inappropriately given the same weight as the medium values *and*, because the distribution of economic values is skewed, this impact is not 'cancelled out' by an inappropriately high weight being given to the low values.

Given the above, we find that the median of the six credible scenarios is likely to be a better estimate of the mean of the population of credible scenarios than is the mean of the six credible scenarios.

2.4. What Can be Learned from the Construction of the Regulatory Test

It is relevant to note that this finding is also consistent with the way the regulatory test is constructed. Part 1(b) of the regulatory test states that in order to pass the regulatory test it is necessary that:

> *'The option maximises the expected net present value of the market benefit … compared with a number of alternative options and timings in a **majority** of reasonable scenarios."*[9] *(emphasis added)*

The regulatory test does **not** state that the option with the highest *mean* net economic value of market benefits passes the regulatory test. By contrast, a necessary (but not sufficient) condition for a project to pass the regulatory test is that the *median* net economic value be positive. That is, a project can not pass the regulatory test unless its sample median economic value is positive.

In doing so, the construction of the test gives low to zero weight to outliers. For example, if there are 6 credible scenarios and an option has slightly negative net value in four scenarios and extremely high value in the remaining two then the option will still fail the regulatory test - even if the mean value is materially positive. This is precisely the situation associated with the DJV investment. In order for the DJV to pass the regulatory test its cost would need to be something less than $136.1m (ie, less than the scenario with the 3rd lowest gross economic value). The AER's proposed adoption of the sample median to determine the DJV's asset base effectively gives the DJV an asset base that is higher than the asset base it would need to pass the regulatory test.

Why the regulatory test is framed in this manner is a highly relevant question in the current context. There are two obvious reasons why the framers of the regulatory test would reject the use of the sample mean in favour of the median (or something lower than the median).

1. The framers recognised that median economic values (generally associated with medium demand forecasts) will be more likely to occur than more extreme economic values. This led to rejection of the mean sample values in favour of the median; and/or
2. The framers considered that, even if each scenario was equally probable, it was nonetheless important that a project only proceed if customers would benefit in at least 50% of the possible states of the world.

---

[9] ACCC, Decision: Review of the regulatory test for New Augmentations, 11 November 2004, page 7.

Rationale one above is consistent with the reasons given in 2.3 above for favouring the sample median even if the ultimate objective is to estimate the population mean.

Rationale two implies that, even if each scenario could be shown to have an equal probability, consistency with the regulatory test would require the AER to adopt the sample median. Consistency is potentially desirable on equity grounds, ie, why should customers pay more for a project than the maximum value at which the regulatory test would have allowed the project to proceed? It may also be desirable because inconsistency would create an incentive to game the regulatory process. For example, proponents of an investment likely to fail the regulatory test but nonetheless with a high mean economic value (due to high outliers) might decide to proceed with investment as an MNSP with an intention of converting to regulated status.

3. Conclusion

We find that the AER's adoption of the sample median is reasonable and that the DJV's submission supporting the adoption of the sample mean is unpersuasive on the following grounds.

1. If the AER is truly, as suggested by the DJV, attempting to estimate the population median then the arguments presented in the DJV submission do not support the use of the sample mean;
2. Consistency with the construction of the regulatory test is likely to be desirable and suggests adoption of the sample median; and
3. The sample mean is likely to be an upward biased estimate of the population mean because high and low demand scenarios are likely to be less probable than the medium demand scenarios and because the relationship between demand and economic value is skewed. (Noting also that the construction of the regulatory test is consistent with such a view regarding relative probabilities of demand scenarios).

NERA
Economic Consulting

NERA Economic Consulting
Level 16
33 Exhibition Street
Melbourne 3000
Tel: +61 3 9623 5020
www.nera.com

NERA Australia Pty Ltd, ABN 34 092 959 665