CENTRE FOR
INTERNATIONAL
ECONOMICS

Review of consumption forecasts

Analysis to support 2006 Bulk Water Price Determination

Prepared for

Independent Pricing and Regulatory Tribunal of NSW

Centre for International Economics
Canberra & Sydney

February 2006

ABOUT THE CIE

The Centre for International Economics is a private economic consultancy operating out of Canberra and Sydney. It undertakes economic analysis for clients around the world.

The CIE solves problems for clients by rigorously analysing markets and regulations, appraising risks and evaluating strategies. We build economic and strategic frameworks to distil complex issues to their essentials. In this way we are able to uncover new insights about emerging developments and assess payoffs from alternative strategies.

The firm has been operating since 1986. Contact details are set out below and more information on what we do and our professional staff can be obtained from our website at www.TheCIE.com.au.

The CIE also co-produces a quarterly report called Economic Scenarios. This analyses global risks and scenarios and can be accessed from www.economic.scenarios.com.

CANBERRA

Centre for International Economics
Ian Potter House, Cnr Marcus Clarke Street & Edinburgh Avenue
Canberra ACT 2601

GPO Box 2203
Canberra ACT Australia 2601

Telephone +61 2 6245 7800 Facsimile +61 2 6245 7888
Email cie@TheCIE.com.au
Website www.TheCIE.com.au

SYDNEY

Centre for International Economics
Suite 1, Level 16, 1 York Street
Sydney NSW 2000

GPO Box 397
Sydney NSW Australia 2001

Telephone +61 2 9250 0800 Facsimile +61 2 9250 0888
Email ciesyd@TheCIE.com.au
Website www.TheCIE.com.au

Contents

1 Introduction

THE INDEPENDENT PRICING AND REGULATORY TRIBUNAL OF NSW (IPART or the Tribunal) provides independent oversight of prices charged by monopoly service providers. Amongst other things, the Tribunal is responsible for setting maximum prices for extraction of bulk water from regulated rivers, unregulated rivers and groundwater. As such, it regulates the bulk water services provided by State Water Corporation (SWC) and Department of Natural Resources (DNR).

The Tribunal is currently revising the price determination for the bulk water services provided by these entities. The price review will establish the maximum charges for bulk water services that apply for a period of five years, starting 1 July 2006.

The determination is intended to set charges at a level that provides the agencies with adequate revenue to ensure reasonable service delivery. To do so, the Tribunal sets prices based, in part, on its estimates of each agency's revenue requirements. This approach requires calculating prices based on assumptions regarding the quantity of water SWC and DNR will sell during each year in the period for which the price determination applies. Given the Tribunal's approach to setting prices, the forecast of consumption is a key input into the price determination.

In the 2006 Bulk Water Price Review, SWC proposed a method for forecasting consumption. This approach reflected forecasting consumption as the long run average usage less one standard deviation. DNR, however, has not commented on forecasting methods. With its proposed a set of fixed charges for water users, DNR's cost recovery would be not sensitive to forecasted consumption.

Terms of reference

Given the sensitivity of the price determination to forecasts of consumption, IPART has contracted the CIE to provide an independent review of the proposed consumption forecast. The primary objective of the review is to assess the reasonableness of the consumption forecast submitted by SWC and DNR for the bulk water review and, if appropriate, to provide the Tribunal with revised forecasts. The scope of this review includes only regulated rivers.

In determining reasonableness, the Tribunal asked that CIE consider four questions. They are:

- Is the approach to consumption forecasting reasonable, for the purposes of setting prices?
- Are the assumptions used by the agencies reasonable and fit for purpose?
- Is the methodology properly applied?
- Is there a balance between use of historical trends and key drivers in generating forecasts?

When formulating responses to these questions, CIE reviewed the agencies' approaches, as well as alternative methodologies. As part of the review, CIE considered the data and assumptions used to generate the forecasts. CIE also considered the implications for forecasting of changing water management rules, particularly those caused by the recently implemented Water Sharing Plans (WSPs).

2 Approaches to forecasting

BULK WATER has several fundamental features that make forecasting consumption challenging and different from that of forecasting urban water consumption. These defining features relate to the following:

- bulk water usage in many of the State's regulated river valleys have frequently been supply-side constrained. The availability of water is necessarily subject to weather conditions/climatic trends; and
- information and data available to inform a bulk water consumption forecast (eg, metered extraction levels) is limited, dating back approximately 20 years.

These factors have a direct impact on the viability and effectiveness of options for modelling bulk water consumption.

This chapter provides an overview of two approaches to forecasting consumption of bulk water. It begins with a discussion of the models and data that exist that can feed into a consumption forecasting exercise. It then discusses the approach proposed by SWC and an alternative time series approach. In discussing the two approaches, the chapter outlines challenges, limitations and strengths of each.

Modelling challenges

Medium to long term forecasting of bulk water consumption necessarily relies on data and information that can assist in predicting supply. (An additional factor is water management rules that can impose a ceiling on extractions.[1]) The availability of water is largely dictated by climatic trends and weather patterns. Therefore, forecasting medium to long term consumption ideally requires reasonable forecasts of future weather patterns. Consultations with DNR and SWC revealed that both agencies

[1] The Water Sharing Plans (WSPs) establish a 'long term extraction limit' for each river valley. However, the limit can be exceeded in any given year under certain circumstances. See Part 8 of any gazetted WSP for further details.

have had limited success in discerning patterns of persistence and climate variability in the data, upon which forecasts can be based.

Demand side factors have, to some extent, limited explanatory power for medium to long term forecasting. For most river valleys, an embargo on new access licences has been in place since the 1980s. While the recently introduced Water Sharing Plans (WSPs) do recognise that long term extraction limits could increase over time in some instances, this increase is capped (generally at around 3 per cent). However, demand side factors (such as area cultivated, crop type and storage practices) can aid with short term forecasting of consumption. They have the potential to inform short term demand forecasts by accounting for such influences as lags and carry overs of water allocations for future rather than current use (ie, up to two years).

IQQM

The existing approach to medium term forecasting of bulk water consumption is based on output from DNR's Integrated Quantity and Quality Model (IQQM) rather than on actual extraction or sales data.

DNR developed IQQM to evaluate the long term impacts of various water management regimes. IQQM does not forecast consumption. Instead as a hydrological model, it captures water availability in a system. This enables estimation of extraction levels that could have occurred over a series of years if the data inputs and scenario definitions were applied in the past. Essentially, the model 'predicts' how a system would have behaved given inputs to flows and storage. The long term data inputs relate to rainfall and gauged flows. Snap shots of the level of 'development' (irrigation and urban extraction drivers) and extraction regulation are imposed on these inputs as a 'scenario'.

The modelled output submitted by SWC and DNR is based on the 'latest scenario.' This scenario is consistent with that which was used in developing WSPs. It reflects development (agricultural and urban), location of extraction points (and total entitlements associated with these points), water user management practices and infrastructure (owned by SWC and major users, eg, irrigators) in place as of 1999–2000, as well as the water management rules that took effect in 2004. IQQM then allows these inputs (or calibrations) to be imposed on historical flow data to model extractions. The output represents potential or modelled extractions given historical flow conditions and current management rules.

Quality of information and data challenges

The feasibility of forecasting and modelling is always constrained by data availability and quality. Information and data available to inform a bulk water consumption forecast is limited. Metering of water extractions did not have wide spread application until the mid 1980s. Furthermore, since then meters are not read with any regularity (or frequency), offering, at best, an annual extraction figure.

Unlike urban water consumption, the information needs relating to determinants of bulk water consumption are not easily met. Demand for bulk water is dependent on a number of factors, such as crop price, land development, on-farm water management practices. This type of information is generally available in varying degrees of completeness and detail. For example, the ABS Agricultural Census, which is administered every five years, collects some of this information. SWC will collect some of this information but not in any consistent format (ie, SWC does not administer regular surveys or store such information).

Actual sales and/or extraction data

An alternative to using IQQM output for forecasting is the use of actual sales or extraction data. The extraction (or usage) data reflects actual metered data. In theory, this metered data is the data set most closely aligned with modelled data. The sales data reflects only the subset of extractions that have been billed. The sales data will not always align with metered data, because not all extractions are billed. In addition, a separate database is used for each data set. As a result, each database is organised around geographical areas and users (ie, licence categories) that are not always consistently defined.

Both SWC and DNR contend that the modelled output is preferable. The modelled data better reflects potential consumption under the current management rules and development conditions. The actual data reflects extraction under the rules and conditions at each historical point in time.

Drawbacks associated with using the actual data (ie, extraction through time) can be summarised to include the following points.

- Actual consumption data is limited. Reasonably reliable data on extractions or volumetric based sales dates back approximately 20 years when metering became standard practice (as part of the conversion to volumetric licences).

- Water management rules have changed substantially in the last 20 years. For example, the change in rules has affected total diversions by up to 12 per cent in some valleys. The WSPs took effect in 2004. The impact of these changes on actual consumption would have to be controlled for in any forecasting exercise based on actual consumption, and this presents difficulties.
- Some valleys underwent significant structural change as a result of both SWC's and irrigators' investment in water infrastructure. For example in the 1980s, the rice industry was largely deregulated, and the cotton industry experienced significant growth. Again, these influences would have to be accommodated.

Using actual data would require controlling for these factors when estimating consumption over the longer term as well as when forecasting consumption under current (or future) conditions.

In the future, actual extraction data may become the preferred data for developing forecasts. However, the case for such a change, in part, relies on the stability of water management rules through time into the future.

SWC's approach to forecasting

To forecast consumption for *pricing purposes*, SWC proposes to use, the long run average of consumption, adjusted downward by one standard deviation, as forecasted consumption. To calculate the long run average (LRA), SWC uses nearly a hundred years of modelled annual extractions, from 1898 to 1993, that are produced by IQQM. This time period is based on a core set of years that is applicable across all regulated river valleys. Many of the regulated river valleys have IQQM output estimates for more recent years, as well as years preceding 1899. This approach effectively reduces the data points available for forecasting consumption. However, the extent to which this biases results depends on the presence of persistence of weather patterns or climatic trends.

The first part of the consumption forecast procedure, ie, to use the long run average, reflects a position held by both DNR and SWC. They contend that the consumption in one year is independent of any previous year(s) and that no meaningful patterns can be discerned from the historical time series on water usage.

The second component of the consumption forecast, ie, to adjust the LRA, SWC describes as a risk adjustment factor. SWC further states that it does not view the (unadjusted) average, mode, or median as appropriate (SWC

2005, p. 123). Consequently, it proposes to adjust this LRA downwards by one standard deviation. SWC justifies this recommended approach as one which is 'commensurate with its increased risk exposure arising from the greater reliance on usage charges, and the water flow variability demonstrated by historical data.' (Submission section 10.3.5) It is important to note that the reduction of the LRA is not because the raw value is regarded as an over estimate of likely consumption. Rather, it is an ad hoc adjustment for risk.

Time series approach

Established methods for forecasting medium to long term bulk water consumption appear to be limited. As noted in the beginning of the chapter, bulk water consumption forecasting is largely confounded by data limitations. Despite these challenges, some options exist. These include a time trend (or structural break) adjustment to the LRA or undertaking a more robust time series analysis.

Trend adjustment

Time trends and structural breaks are just two examples of myriad of issues affecting considerations of alternative approaches to forecasting consumption. For example, do climatic (and therefore hydrological) conditions seem to have changed over the past century? In technical terms, are there 'trends' or 'structural breaks' in the data series on water usage? Using the historical mean may be inappropriate if water usage in a river system is trending up or down over time, or if there are climatic cycles.

The following example illustrates the importance of understanding the climatic conditions that will prevail during the forecasted period and that such trends can potentially affect the forecast. The mean annual modelled water usage in the Macquarie River between 1892 and 1944 (as generated by the IQQM model) was 359 192 ML. The mean for the 1945–95 period was approximately 18 per cent higher (422 866 ML). The modelled 'historical average' over the entire period was 391 883 ML. If one had wished in 1995 to forecast water usage for the next five years (which turned out to average 435 456 ML annually), then clearly the mean for the 1945–95 period would have performed better than the 'historical average'.

Discussions with DNR officials indicate that empirical evidence supports the view that climatic conditions in the first half of the 20th century differed considerably from that of the post-World War II decades. However, both

SWC and DNR suggest that it is not evident whether the weather conditions of the next five years will be a continuation or deviation from what was experienced in the last half of the twentieth century. This example further illustrates the challenges to identifying climatic trends and other factors that are key considerations/inputs to developing reasonable forecasts of bulk water consumption.

Time series analysis

Time series models are used for forecasting future values of variables of interest when only historical data on these variables is available, and there are no structural models available that can explain the behaviour of such variables in terms of that of other underlying variables. In other words, time series econometrics is concerned with the estimation of difference equations containing stochastic components. Uncovering the dynamic path of the variable of interest — its time series — improves forecasts since the predictable components of the series can be extrapolated into the future.

In the case of forecasting consumption, using a time series approach can be thought of as a process driven ultimately by annual changes in weather conditions. It utilises statistical techniques to extract additional information from the available data series on modelled water usage to improve the forecast of future water usage.

CIE opted to employ an autoregressive integrated moving average (ARIMA) approach as an alternative method to forecast consumption. This approach discerns some pattern in consumption from the modelled historical data, and postulates that the pattern is based on some statistical correlation (relationship) between current and past consumption. The premise of the ARIMA model is distinct from the LRA, which assumes that consumption in any given year is independent of the last.

To understand what an ARIMA model is, it is helpful to first understand two classes of simpler time series models, the AR models and the MA models.

AR and MA time series models

Time series models have autoregressive and/or moving average components. Autoregressive (AR) time series models consist of past observations of the dependent variable (ie, the variable of interest) in the forecast of future observations. For example, the simplest AR model, the AR(1) model, includes one lag of the dependent variable:

$$y_t = a\, y_{t-1} + e_t,$$

where y is the dependent variable, a is a parameter, and e_t is the random error or white noise term. We can think of e_t as the forecast error.

Moving average (MA) models include past observations of the white-noise process (that is, past forecast errors) in the forecast of future observations of the dependent variable. Technically speaking, a sequence $\{e_t\}$ is a white-noise process if each value in the sequence has a mean of zero, a constant variance, and is uncorrelated with all other realisations in the sequence. The MA(1) model includes one lagged observation of the white-noise process:

$$y_t = b\, e_{t-1} + e_t,$$

where e_{t-1} is the lagged observation of the noise process.

Autoregressive Moving Average (ARMA) models comprise both past observations of the dependent variable and past observations of the innovations noise process in the forecast of future observations of the dependent variable of interest. For example, the ARMA (2,1) model may be represented as:

$$y_t = a_1\, y_{t-1} + a_2\, y_{t-1} + b_1\, e_{t-1} + e_t.$$

The generic ARMA (p, q) model contains p number of past observations of the dependent variable and q number of past observations of the white-noise process.

By definition, an ARMA model is covariance stationary in that it has a finite and time-invariant mean and covariances. Shocks to a stationary time series are necessarily temporary; over time, the effects of the shocks will dissipate and the series will revert to its long run mean level. When the time series is not stationary, it may be necessary to remove the trend by repeated differencing. For example, the d-th difference of a generic Autoregressive Integrated Moving Average ARIMA (p, d, q) model is stationary. That is, an ARMA model is an ARIMA model which is stationary.

Implementing an ARIMA model

Operationalising the ARIMA model involves finding suitable values for the p, d and q parameters. The Box-Jenkins (1976) strategy is commonly used to identify the most appropriate specification for the ARIMA model. The strategy consists of 3 stages, and can be implemented in many of the econometrics (regression) software packages available. CIE used SHAZAM

Professional Edition, released by Northwest Econometrics of Vancouver, Canada.

The following details the three-stage process.

Stage 1: Model identification

The first step is for the researcher to visually examine the time plot of the historical data series, the autocorrelation function, and the partial correlation function. Plotting the time path provides useful information concerning outliers, missing values and structural breaks in the data. The autocorrelation function (ACF) is a correlation sequence of a random time series with itself, while the partial correlation function (PCF) is a correlation sequence estimated by fitting successive order autoregressive models to a random time series by least squares. The rates of decay of the ACF and PCF will give an indication of the stationarity of the model and inform the statistician whether further differencing is necessary. If the model appears stationary, plausible values for p and q in the ARIMA (p, d, q) model can be inferred from the patterns found in plots of the ACF and PCF.

Stage 2: Model estimation

In this stage, each of the tentative or plausible models is fitted and the various a_i and b_i coefficients examined. Box and Jenkins believe that parsimonious models (those with small values for p and q) produce better forecasts than overparameterised models. The aim is to approximate the true data-generating process but not to pin down the exact process. Two of the more commonly used model selection criteria are the Akaike Information Criterion (AIC) and the Schwartz Bayesian Criterion (SBC). Ideally, the AIC and SBC should be as small as possible. In addition, the t-statistics for all the a_i and b_i coefficients should also be statistically significant and the adjusted R^2 statistic for the fitted regression reasonably high. (Sometimes, human judgement based on experience is required when there is trade-off between the various criteria when choosing between two candidate models.) Further diagnostic checking can also be implemented, such as plotting the residuals of outliers and looking for evidence of periods in which the model does not fit the data well.

Stage 3: Forecasting

In SHAZAM, the forecasting procedure is made operational by using the model parameter estimates and the estimate of the error variance from the estimation stage. If a moving average component is present then estimated residuals also enter the forecast. Consequently, a sequence of one-step ahead prediction errors is computed to the end of the sample period (including back-forecasting of pre-sample residuals). Point forecasts are calculated recursively from the difference-equation form of the process.

The ARIMA forecasts are compared with the historical averages used in State Water's submission to IPART and augmented using DNR's modelled IQQM output for more recent years where available (ie, 1993 to 2002). Forecast performance is measured by summing the absolute (or squared) deviations of the ARIMA annual forecasts from the IQQM figures over a given five-year period.

2.1 **Implementing the ARIMA model for the Border Rivers**

Model identification stage

'Historical data' on IQQM modelled output from 1898 and 1997 were used as input into the ARIMA model.

Setting d = 0 initially, plotting the ACF and PCF showed a failure of both functions to converge or decay sufficiently even after 100 periods, indicating the presence of a trend in the historical data (and the lack of stationarity). When d was set to 2, the ACF and PCF plots looked more satisfactory, suggesting that stationarity was achieved.

Model estimation stage

Having determined the value of d, the next stage involved finding the most appropriate combination of values for p and q. After trying different permutations of p and q values, it was found that the ARIMA model with p = 1 and q =1 yielded the lowest (and therefore the best) value for the Schwartz criterion, which was 21.957. The adjusted R-square statistic for the regression was 0.6949 (indicating good overall fit) and the t-statistics for the AR and MA parameters were –4.514 and 124.2, which indicated very high levels of statistical significance.

Forecasting stage

The ARIMA(1,2,1) model with 1898-1997 data was used to forecast water usage from 1998 to 2002.

ARIMA modelling results

Table 2.2 summarises the ARIMA model used for each regulated river valley. Recall that p indicates the number of autoregressive lags (past water usage figures), d indicates the number of times the data needed to be differenced in order to be de-trended, and q indicates the number of historical forecast errors that are useful in forecasting future water usage.

2.2 ARIMA model used for each river valley

Regulated river valley	*ARIMA model parameters*		
	p	***d***	***q***
Border	2	1	1
Gwydir	1	1	1
Hunter	1	1	2
Lachlan	0	1	2
Macquarie	0	1	1
Murray	0	1	1
Murrumbidgee	0	1	1
Namoi	1	2	1
Peel	1	1	1

Source: CIE.

3 Assessing forecasting approaches

THIS CHAPTER assesses the relative merits of two approaches to forecasting consumption, the SWC's proposed approach and the alternative ARIMA approach. As noted previously, the SWC's proposed approach has two components that can each be analysed independently. The first is forecasting consumption using the long run average (LRA). The second is adjusting the long run average downwards by one standard deviation. The chapter begins with a discussion of the adjustment. It then presents the results of comparing forecasting consumption using a time series approach to using the long run average.

Implications of adjusting the mean by one standard deviation

If usage were constant over any time period, a volumetric charge (per ML) could be calculated that would exactly recover the cost component assigned to usage by dividing that cost (C) by the usage (V). With varying usage, a forecast of V is required if a constant price is adopted. Under and over recovery of costs in any one year depend on the divergence between actual and forecast V. SWC has proposed forecasting the usage by the modelled long term average V but then reducing that forecast by one standard deviation (calculated from the modelled time series that produced the long term average). This results in a new 'forecast' volume which is a deliberately downward biased constant for the ensuing determination period. The amount of bias increases with the degree of variability (as measured by the standard deviation σ) about the long term average in each valley.

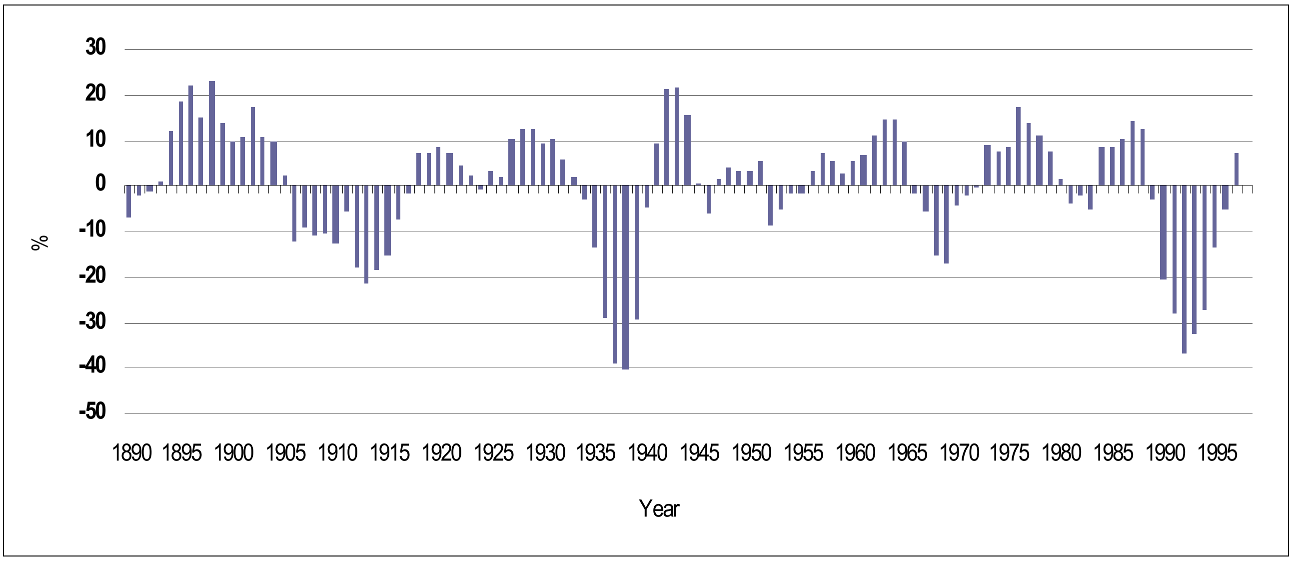
This means that there will be an *expected* over recovery of costs for the next five-year period. SWC has suggested that an adjustment mechanism would then be used to reduce the required revenue and hence the price that would otherwise be set in the future determination period to compensate for this over recovery. (Under recovery, if it occurs, would be followed by a subsequent upward adjustment to price for the next five years). However, if the (unadjusted) long term average is the best unbiased estimate of consumption for any arbitrary five-year period, this approach will produce an average tendency to over recover with delayed compensation to users. SWC seeks to justify this approach on the grounds of the increased risk they face through the move to greater reliance on usage pricing as a contribution to cost recovery and the level of variability which varies between valleys.

It turns out that the SWC's approach of using a standard deviation to ameliorate revenue risks is extremely conservative. In most five-year periods, it is almost certain to produce a large over-recovery of costs and potentially trigger a substantial downward price adjustment in the next period. This can clearly be seen by comparing charts 3.1 and 3.2. (Appendix A provides similar charts for eight other regulated river valleys.) Chart 3.1 shows the aggregated five-year (forward) deviations from mean modelled water usage in Macquarie river from 1890 to 1996. For example, the 1890 data point shows (in percentage terms) the sum of the differences between IQQM modelled water usage and the 'historical mean' (ie, LRA) for the years 1890, 1891, 1892, 1893 and 1894. Chart 3.2 shows these deviations from the 'historical mean' minus a standard deviation for the same river and period. It is obvious that in virtually every five-year period (except 2 out of 107), SWC's revenues would have exceeded costs. Moreover, in a vast majority of these periods, the average deviation from the historical mean would have exceeded 10 per cent, thereby triggering the proposed ex-post adjustment mechanism and generating price instability between five-year periods.

This approach begs the question of why a more orthodox approach to dealing with the cost of risk would not be better — one which did not rely on systematically biasing the usage forecast by greater or lesser amounts in different valleys but which could still rely on the information on variability in each valley's time series of usage.

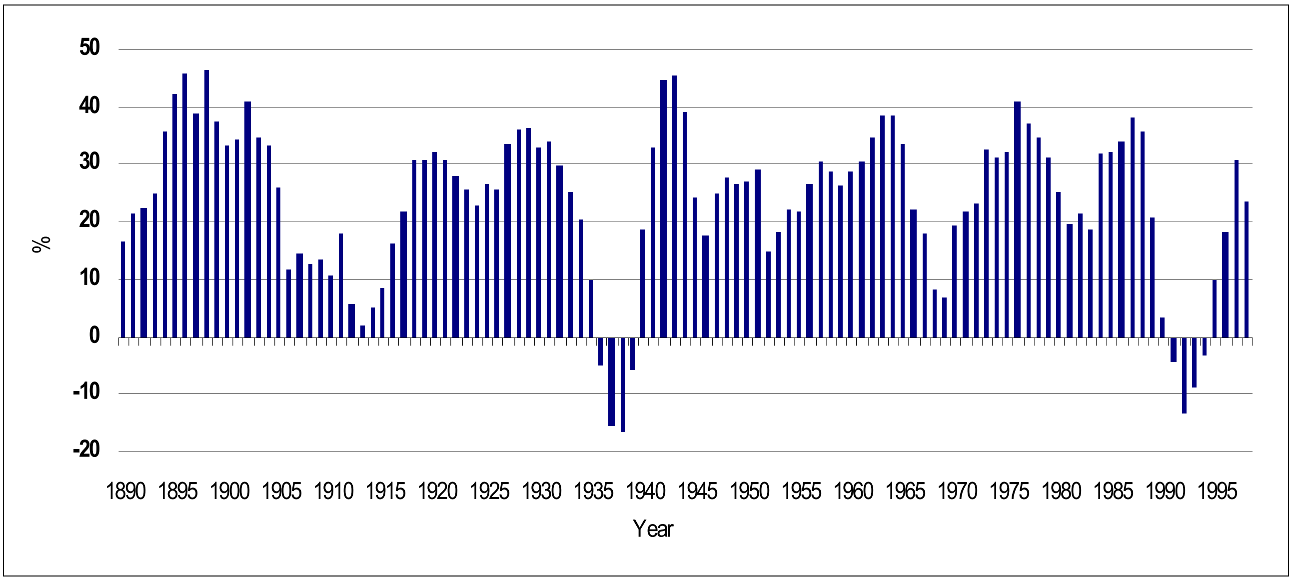
For instance, if it is the overall revenue variability that matters to SWC as a business, this could be captured in recommendation of a higher weighted average cost of capital (WACC), the effects of which could then be distributed across the cost recovery burden for individual valleys to reflect their relative contributions to revenue instability. SWC could then provide estimates of the costs of hedging against increased exposure to revenue risks in the financial markets. This would make the 'costs of risk' just another explicit cost which needed to be recovered and raise the transparency of SWC's cost recovery formula. It would be relevant to demonstrate how differently this might impact the relative prices in the different valleys compared with the SWC 'adjusted forecast' approach.

3.1 **Aggregated five-year (forward) deviations from mean modelled water usage (Border rivers)**



Data sources: DNR; CIE.

3.2 **Aggregated five-year (forward) deviations from mean minus 1 standard deviation of modelled water usage (Border rivers)**



Data sources: DNR; CIE.

Performance assessment

To assess the potential gain of using a more robust time series approach to forecasting, CIE applied an ARIMA model to nine regulated river valleys, forecasting annual consumption over a five-year period. CIE then compared the ARIMA results to the (unadjusted) long run average.

The assessment of each approach relied on three main measures. They are bias, accuracy and goodness of fit. Bias considers whether the approach systematically over- or under-estimates consumption. Accuracy is measured by how close the forecasted estimate is to the benchmark (ie, the magnitude of the difference between the forecast and the benchmark). Goodness of fit looks at how well the pattern of consumption is forecasted relative to the benchmark. With each of these measures, the assessment considers the robustness of the approach. The robustness looks at the overall performance of the forecast approach, considering all nine rather than any one regulated river valley.

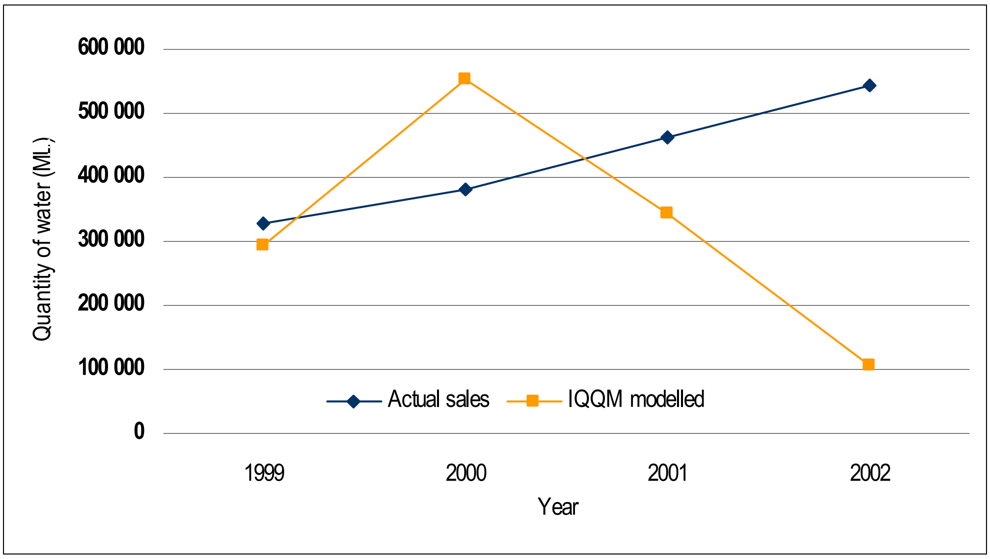
Benchmark for assessment

As noted, determining the effectiveness and robustness of the two forecasting approaches requires a benchmark for evaluation. This assessment uses the IQQM modelled extractions as the benchmark. In other words, it assesses how forecasted consumption in any given year (based on ARIMA and LRA approaches) compares to IQQM modelled extraction for that same year.

The benchmark, while pragmatic, has some limitations. To an extent it may provide a 'false comparison' in that the benchmark reflects modelled rather than actual extractions. However, using actual metered data (rather than modelled extractions) is not appropriate.

A comparison of actual extractions to IQQM output shows how the two can deviate substantially. The levels and changes in extractions do not always move together. The IQQM output imposes, retrospectively, WSP rules. The difference between IQQM and actual extractions in these years highlights, in part, the extent to which water management rules, development and other factors can influence actual consumption. This point is particularly evident when comparing the years directly preceding the implementation of WSPs. The graph below (chart 3.3) illustrates these points.

3.3 **Actual sales versus modelled IQQM output, Macquarie river, 1999-2004**



Data sources: DNR; NSW State Water.

ARIMA and LRA estimates

The table below provides a summary of the five-year average consumption forecasted using the ARIMA and LRA approaches, as well as the benchmark value derived from IQQM output.

3.4 **Average consumption[a]**

Regulated river valley	***Forecast period***	***Benchmark***	***LRA***	***ARIMA***
		ML	ML	ML
Border	1998–2002	208 240	290 670	222 240
Gwydir	1997–2001	403 110	309 160	359 000
Hunter	1990–94	121 540	128 070	130 840
Lachlan	1995–99	343 550	307 150	340 020
Macquarie	1996–2000	435 460	386 310	434 440
Murray	1995–99	2 208 950	1 934 830	2 136 170
Murrumbidgee	1998–2002	1 933 640	1 915 850	2 080 550
Namoi	1997–2001	238 980	237 150	221 430
Peel	1995–99	13 770	14 680	15 050

[a] Benchmark and ARIMA estimates represent the five-year average during the forecast period.

Source: CIE.

Rather than looking at a consistent period of time across each river valley, CIE opted to maximise use of the available data. The period of time for which there is IQQM modelled output varies across river valleys. For example, the Hunter river valley only has IQQM output to 1994, while the Border and Murrumbidgee river valleys have IQQM output up to 2002. As a result, the forecast period varies across river valleys.

The period of time from which SWC estimates the LRA and applies this estimate as a forecast is different from that of the benchmark and ARIMA values. SWC's LRA estimate is calculated using IQQM output from 1893 to 1993. It then uses this figure as the estimated annual consumption in each year of the forecast period. ARIMA uses IQQM output from 1893 to the mid 1990's to estimate annual consumption in five-year forecast period. Unlike the LRA, the ARIMA consumption estimate varies from year to year in the forecast period.

Bias and average accuracy

Bias considers whether either approach systematically over- or under-estimates consumption, while average accuracy measures how close the average estimate is to the benchmark (ie, the magnitude of the difference between the five-year average forecast and the benchmark). The benchmark here is defined as the five-year average for the forecast period using IQQM output. Table 3.5 summarises the performance of each approach as indicated by these two measures.

Generally both approaches to forecasting consumption have the same bias. In only one river valley, the Murrumbidgee, the estimates do not share the same bias. The ARIMA over estimates (modelled) consumption, while the LRA underestimates consumption. In the other eight river valleys, both the ARIMA forecasted five-year average and the LRA tend to have the same bias.

Neither approach appears to have any systematic bias. ARIMA over estimates consumption in four of the nine river valleys. LRA over estimates consumption in three of the nine. Also, over- and under-estimating consumption does not appear to be associated with geographic location. For example, ARIMA underestimates consumption for the Murray river valley, while it overestimates consumption in the Murrumbidgee river valley. The northern river valleys (eg, Border, Gwydir, Namoi and Peel) show similar diversity in over and underestimating consumption.

While both approaches, LRA and ARIMA, tend to show the same patterns of bias, ARIMA appears to be slightly more accurate. In five of the nine river valleys, the five-year average consumption forecasted using the ARIMA approach is closer to the benchmark (ie, five-year average for the forecast period using IQQM output). In addition, when taking the average across the river valleys, the difference between the benchmark and ARIMA values tends to be about 6 per cent.

The LRA average tends to perform slightly worse than the ARIMA model. Across the river valleys, the difference between the benchmark and LRA tends to be about 12 per cent.

Goodness of fit

Goodness of fit looks at how well the pattern of consumption is forecasted relative to the benchmark. This measure of effectiveness considers how well each approach performs in any given year, rather than comparing the five-year average derived from the forecasted annual consumption.

To measure goodness of fit, the analysis uses the sum of the absolute deviations in each year of the forecast period. Table 3.5 presents the sum of the absolute deviations for the ARIMA and LRA approaches.

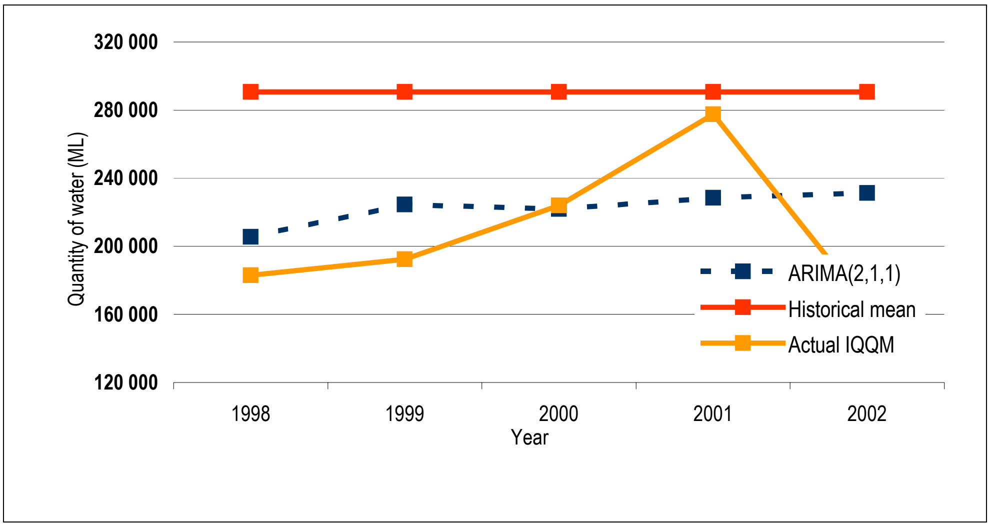
3.5 **Measure of performance — bias, accuracy and goodness of fit**

	Bias & accuracy % difference from benchmark		***Goodness of fit sum of absolute deviations***	
	ARIMA	LRA	ARIMA	LRA
Border	7	40	172 091	412 141
Gwydir	-11	-23	383 515	492 400
Hunter	8	5	96 281	100 590
Lachlan	-1	-11	291 536	344 325
Macquarie	-0.2	-11	702 860	750 990
Murray	-3	-12	628 210	1 370 598
Murrumbidgee	8	-1	1 057 246	898 444
Namoi	-7	-1	113 965	119 075
Peel	9	7	7 274	6 496

Source: CIE.

ARIMA generally has a better fit when compared with the LRA. In seven of the nine river valleys, the sum of the absolute deviations for the ARIMA approach is less than that of the LRA (the same is true for the sum of squared deviations). This result suggests that in any given year, the forecasted consumption using an ARIMA approach tends to be closer to the benchmark. Chart 3.6 shows the goodness of fit of the forecasts generated by the two alternative approaches for the Border rivers between 1998 and 2002.

3.6 **ARIMA forecast against LRA and IQQM output: Border Rivers (1998–2002)**



Data sources: DNR; CIE.

4 Conclusions

THIS CHAPTER discusses the conclusions that can be drawn from the analysis.

The use of consumption as modelled based on DNR's IQQM is reasonable given the available data and absence of an alternative model for forecasting.

IQQM has undergone independent assessments and audits. While it is not designed to forecast, it is a reasonable means for controlling a range of influencing factors (eg, development, changes in water management rules). Use of actual extraction and/or sales data would necessarily truncate the data available upon which forecast can be based, limiting the precision of the forecast. In addition, using actual data would require controlling for factors that are addressed in the latest scenario used in IQQM.

At this point in time, using actual metered data (rather than modelled extractions) is not appropriate for several reasons. Metered data is limited, dating back roughly 20 years. Water management rules have undergone significant changes as a result of recently implemented WSPs, reducing the relevance of past actual consumption behaviour for forecasting future usage. Lastly, drivers of demand for bulk water have changed through time, such as area cultivated, crop type, infrastructure and water management practices. The impact of these changes on actual consumption are not easy to control for, given limited data on each influence.

In the future, as the data points for actual extractions become larger and if water management rules (and other factors) stabilise, actual metered data may become a more appropriate basis for forecasting consumption.

ARIMA better utilises the information from the IQQM data series. However, both ARIMA and LRA approaches do not forecast well for periods with significant 'turning points'.

Both the LRA approach and the ARIMA approach utilise historical data. However, they employ this data differently when forecasting consumption. The LRA approach implicitly assumes that consumption in any given year is independent of the previous year. Consequently, the SWC's approach results in consumption over a five-year forecast period to be estimated as constant and equivalent to the historical long run average.

From a forecasting or modelling perspective, ARIMA better utilises the information from the IQQM data series. ARIMA attempts to discern patterns of consumption from the historical data and, in turn, uses these patterns to forecast consumption. The result is that consumption in any given year of the forecast period can vary. However, ARIMA also attempts to forecast consumption while 'smoothing' the change in consumption from one year to the next. (See Chart 3.6 to illustrate how ARIMA and LRA change from year to year in a five-year forecast period.)

The both approaches tend to 'smooth' consumption, ie minimise change in consumption during the forecast period. (LRA does this simply by using a constant value for the forecast period.) This tendency results in the models having very limited ability to forecast periods with significant turning points, as seen at the start of severe drought periods (eg, 2002 to 2004).

An ARIMA approach generally performs slightly better than the LRA when forecasting consumption.

When assessed relative to the LRA approach, ARIMA performs better in two of the three performance criteria. As discussed, ARIMA and LRA generally exhibit the same bias. In eight of the nine regulated river valleys that were modelled, ARIMA and LRA approaches had the same bias.

ARIMA, however, performed slightly better with respect to accuracy. In five of the nine river valleys, ARIMA was more effective at forecasting a five-year average of consumption relative to the LRA. In any given river valley, the ARIMA forecasted five-year average consumption differed from the benchmark by 11 per cent and averaged, across all river valleys, to be roughly 6 per cent. The largest difference between the benchmark and the LRA was 40 per cent and averaged across the river valley to be 12 per cent. In addition, the sum of absolute deviations is smaller for the time series model in seven of the nine river valleys. That is, the ARIMA model exhibits superior goodness of fit for year by year forecasts in seven out of nine cases.

The accuracy gains of using a time series approach, such as ARIMA, may be limited.

The gains of from an ARIMA approach may be limited when developing a five-year forecast of consumption to be used in price determinations. Bias and average accuracy rather than year by year goodness of fit have the most relevance for how the forecast is 'operationalised' or employed by IPART in price determinations. (The reason is that IPART would presumably like to set a smooth price path for the five-year period rather than one that tracks year by year fluctuations in anticipated consumption.) For both measures, the relatively better performance of ARIMA is minor.

SWC's proposal to adjust the LRA downwards by one standard deviation is inappropriate.

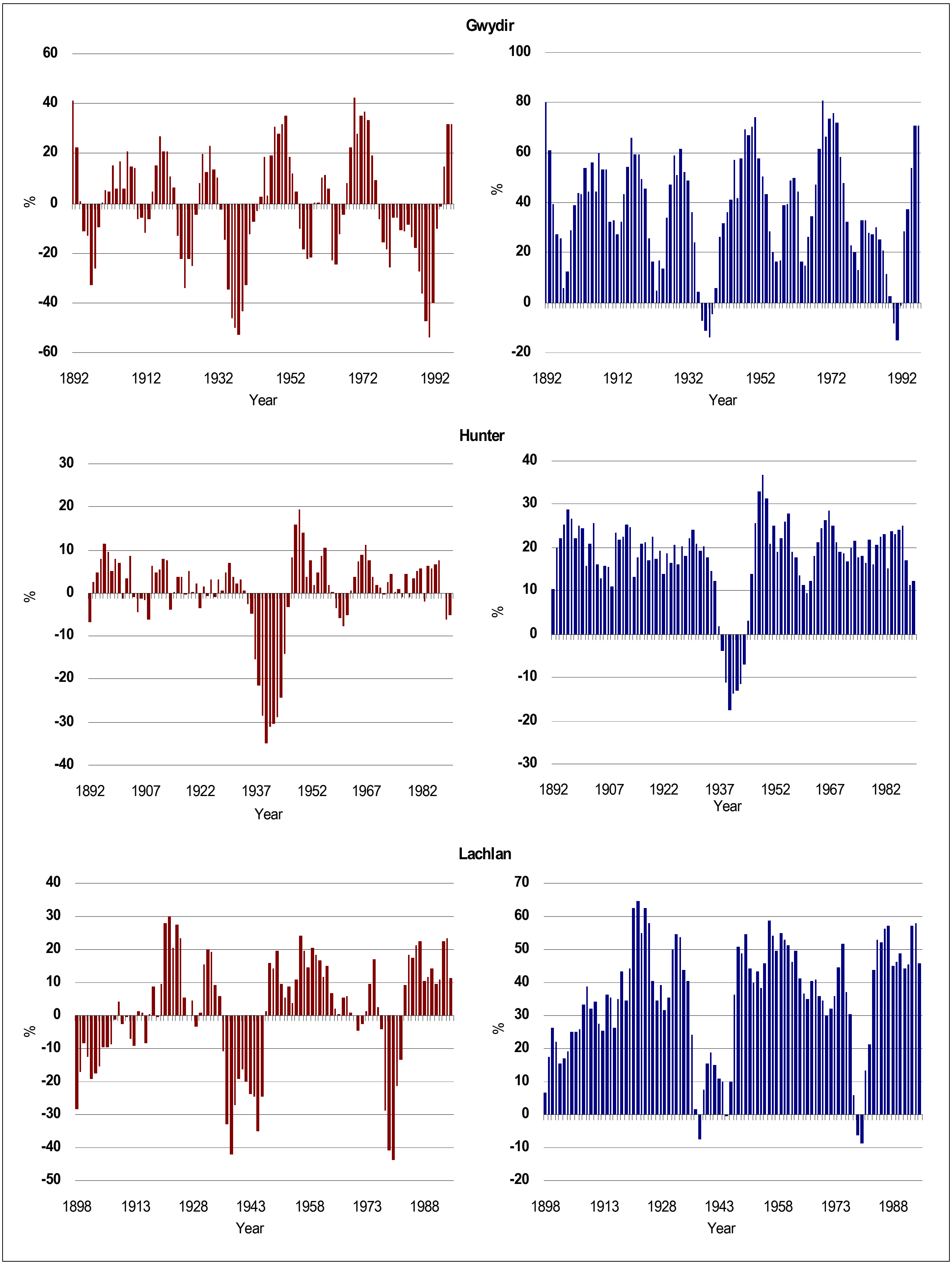
The analysis shows that there is a high probability that SWC's proposed adjustment would result in substantial over-recovery of costs over most five-year periods. Such a result would likely lead to ex-post price adjustments and facilitate pricing instability between periods.

Further, such an adjustment for risk would be better addressed in other elements of the price determination. SWC's approach to managing their exposure to increased revenue risks appears arbitrary and unnecessarily conservative. The costs of hedging in the financial markets against the increased revenue risks arising from the move towards 'user pays' should be made explicit and transparent. The increased costs could potentially be recovered through a higher weighted average cost of capital (WACC) rather than by adjusting the consumption forecast downwards by one standard deviation from the historical data series. If thought appropriate, these costs could then be distributed for recovery across the valleys to reflect the relative contribution of each to the revenue risks faced.

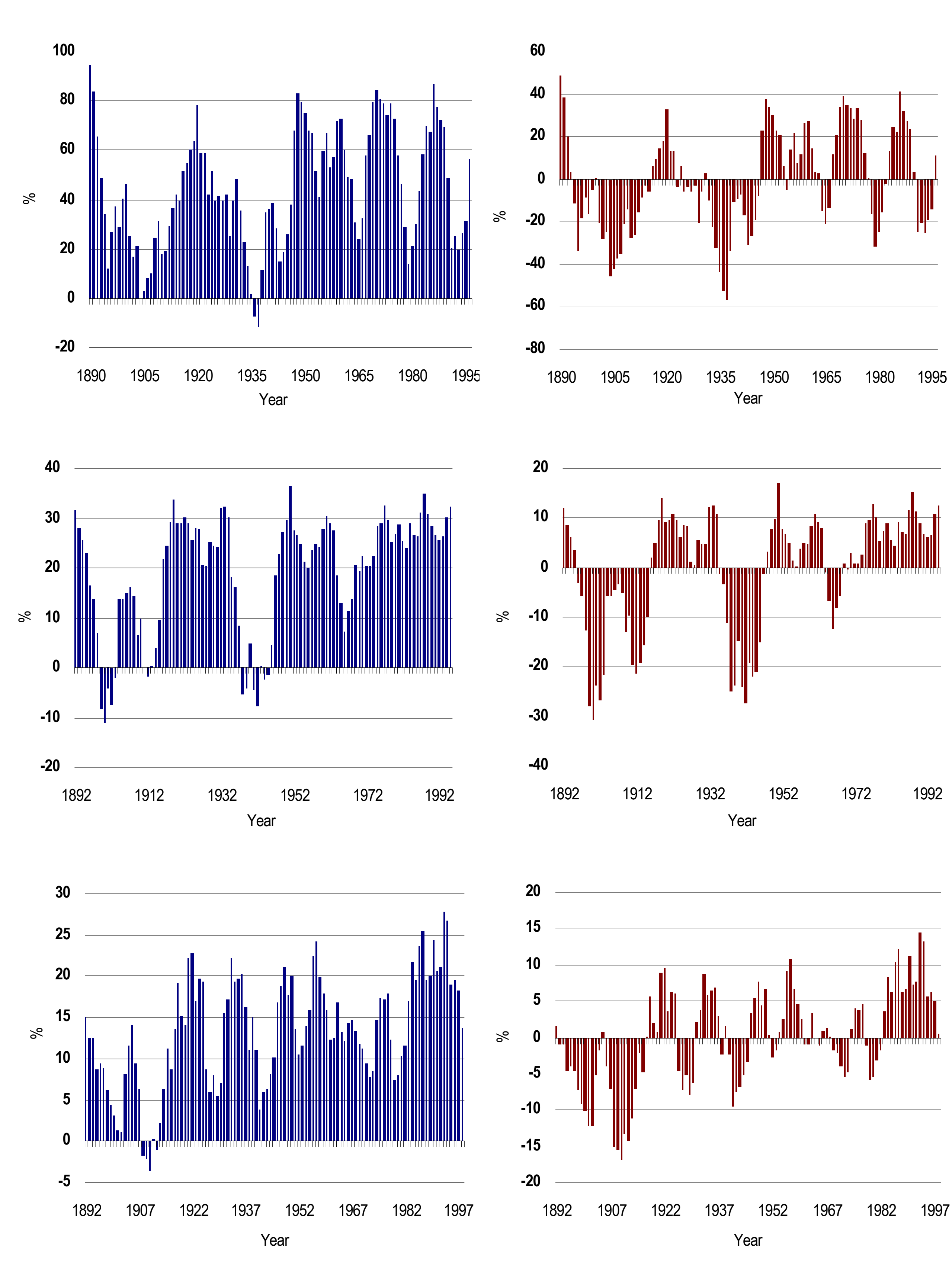
A Aggregated five-year (forward) deviations

THE FOLLOWING provide charts for eight regulated river valleys that illustrate the implications of adjusting the LRA downwards by one standard deviation.

A.1 **Aggregated five-year (forward) deviations from mean and from mean minus 1 modelled water usage — Gwydir, Hunter and Lachlan rivers**



A.2 **Aggregated five-year (forward) deviations from mean and from mean minus 1 modelled water usage — Macquarie, Murray LD and Murrumbidgee rivers**



A.3 **Aggregated five-year (forward) deviations from mean and from mean minus 1 modelled water usage — Namoi and Peel rivers**