Economic withholding approach, limitations and results

Wholesale electricity market performance report 2022

December 2022

AUSTRALIAN ENERGY REGULATOR

Inquiries about this publication should be addressed to:

Australian Energy Regulator
GPO Box 3131
Canberra ACT 2601
Tel: 1300 585 165

AER reference: 14733467

Amendment record

Version	Date	Pages
1	14 December 2022	32

Contents

1 Introduction

To inform the AER's assessment of wholesale electricity market performance, one factor we must consider is whether in the long run wholesale prices are determined by underlying costs. In chapter 6 of our *Wholesale electricity market performance report 2022* (the performance report) we used new metrics to assess potential economic withholding – that is, participant conduct that involves pricing capacity above marginal cost with the intention of influencing the price. These metrics are an initial step towards extending our analysis to screen for potentially harmful behaviour. They were considered alongside a broad range of other factors in assessing the effectiveness of competition in the market.

Since we have limited access to direct information on participant costs, we developed techniques that do not require this. We screened for periods when the price was higher than we might expect given the level of surplus capacity in the market, and we examined generator behaviour in these periods. We also estimated generator incentives to withhold in each period and examined conduct in periods of potentially higher incentive.

This technical paper is complementary to the analysis in the performance report. While the performance report contains the high-level results of the analysis and calculations performed, this technical paper explains the method underlying the calculations and assessment and provides more detailed results. This paper discusses the metrics used to assess *economic* withholding only; metrics used to assess *physical* withholding are outlined in our general methodology.[1]

1.1 Background

The National Electricity Law (NEL) requires the AER to monitor the wholesale market and report on its performance at least every 2 years.[2] We are required to identify and analyse whether:

- there is 'effective competition' within the relevant wholesale market, as defined in the NEL
- there are features of the market that may be detrimental to effective competition within the market
- there are features of the market that may be impacting detrimentally on the efficient functioning of the market and the achievement of the national electricity objective.

Further, the NEL stipulates that our performance report must contain a discussion and analysis of the monitoring methodology applied and the results of indicators, tests and calculations performed.[3]

This methodology document contains a discussion of the economic withholding analysis undertaken in our performance report. We aim to communicate how we developed and implemented our metrics, the data sources we used, and some additional results.

[1] AER, *Wholesale electricity market performance report 2022–Methods and assumptions*, December 2022.

[2] National Electricity Law, Part 3 Division 1A

[3] NEL, Section 18C(3)(d)

2 Summary and overview

This document sets out the methodology underlying the various metrics and analyses used to assess whether there is evidence suggestive of systematic and sustained economic withholding in the National Energy Market (NEM). For this analysis, 'systematic' economic withholding means that at times when a participant has a greater profit incentive to withhold and when it has ability to influence the price, it typically submits higher bids.[4] To evaluate this, we aim to assess whether generators consistently increase their offers for specific periods of time and for reasons unrelated to supply-side drivers (such as input prices, start-up costs and fuel scarcity). Moreover, to test whether withholding behaviour is 'sustained' we observe whether it occurs consistently over several years (which suggests market power is not being competed away).

The most direct way to assess economic withholding behaviour is to compare bids to underlying costs. In the absence of reliable cost information, our approach combined the use of 3 metrics (Table 1) to generate 2 withholding analyses (Table 2).

Table 1: Withholding metrics

Metric	Purpose	Description
Surplus capacity screen	Screen	Identifies "outlier" periods when the dispatch price significantly exceeds the expected dispatch price given market conditions.
Returns to withholding capacity (RWC)	Incentive	Quantifies a participant's incentive to withhold by estimating the additional profit from withholding.
Quantity weighted offer price (QWOP) increase	Measure	Measures how a generator changes its bids relative to typical bidding for similar conditions.

Table 2: Withholding analyses (using the metrics from Table 1)[5]

Metric	Purpose	Description
RWC + QWOP increase	Incentive + Action	Assess the extent that a participant submits higher

[4] This is distinct from a plant economically withholding by submitting higher than competitive bids for all periods, which is beyond the scope of this analysis as it requires information on participant costs.

[5] An extension for future work could consider combining all three metrics – that is, combining the screen (surplus capacity), incentive (RWC) and action (QWOP increase). This approach would assess how a plant changes its bids relative to its typical bidding during outlier periods of greater incentive. We leave this analysis for future work as our current method for estimating incentives is likely to be less robust during outlier periods.

		bids when it has a greater profit incentive to do so.
Surplus capacity screen + QWOP increase	Screen + Action	Identify generators that consistently submit higher bids during outlier periods (i.e., high price periods).

The rest of the document is set out as follows:

- section 3 provides further background on the metrics and analyses, outlining the purpose of each metric and reference material for how we developed it
- sections 4, 5 and 6 set out an overview for each of the 3 metrics, outlining the implementation process, the data requirements and some possible variations on our method
- sections 7 and 8 outline how we combine the metrics to consider two analyses
- the appendices provide further detail on the data, specifically:
  - Appendix A sets out the share of wind and solar generation by region and year
  - Appendix B presents the annual surplus capacity-price relationship and screen for each state
  - Appendix C shows the effect of different interconnector assumptions on the surplus capacity-price relationship
  - Appendix D outlines our method for calculating total effective capacity in a region
  - Appendix E presents selected summary statistics that may assist in replicating our results.

3 Background on metrics and analyses

In this section, we provide additional background on our metrics and then explain our approach of combining the metrics to form the analyses.

3.1 Surplus capacity (screen)

Surplus capacity (also known as 'the supply cushion') measures market tightness by subtracting the current load in a period from the available capacity in that period. That is, surplus capacity is the quantity of generation that can be called upon. When the level of surplus capacity is low, this should typically mean that prices are high. This is because more expensive generators must be dispatched to satisfy the load, therefore leading to a higher dispatch price. The opposite is true for a high level of surplus capacity.

The surplus capacity metric aims to be a market-wide screen for periods in which the dispatch price is significantly higher than would be expected given the tightness of the market. This means it can identify 'outlier' (high price) periods to then focus further analysis on whether the period may be caused by a participant or plant that is withholding.

The surplus capacity (or supply cushion) methodology was developed by the Alberta Market Surveillance Authority, where it is used extensively.[6] The method is also used in Ontario, Singapore, and the Philippines.[7]

3.2 Returns from withholding capacity (incentive)

The returns from withholding capacity (RWC) methodology estimates a participant's profit gain from withholding capacity for a specific settlement period. It attempts to capture the additional revenues a participant could earn by withholding capacity. These additional revenues are a function of the sensitivity of prices to withholding and the amount of remaining operating capacity owned by the participant which would earn higher revenues. This metric only measures the potential gains rather than actual behaviour.

The RWC methodology was developed by academics working for a German think tank called the Monopolies Commission.[8] The Federal Cartel Office, the German competition regulator, has adopted parts of the methodology to conduct various analyses. In an Australian context, the Victorian Energy Policy Center (VEPC) applied the RWC methodology to show that the closure of Hazelwood increased the returns to withholding in the NEM.[9]

---

[6] For methodology, see: Alberta Market Surveillance Administrator, *Supply Cushion Methodology and Detection of Events of Interest, Directions Paper*, June 2012. For applications, see *Quarterly Report for Q4 2021,* February 2022, p. 18 and *Quarterly Report* (*Q4/21),* February 2012, section 2.1, p. 8.

[7] Ontario: Ontario Energy Board, *Monitoring Report on the IESO-Administered Electricity Markets,* January 2014, section 4.2, p. 63. Singapore: Energy Market Authority, *Energy Market Surveillance and Regulation in Singapore*. March 2018, slide 7. Philippines: Philippine Electric Market Corporation, *Annual Market Assessment Report*, August 2020, Annex C p. 55.

[8] Bataille, Bodnar, Steinmetz, and Thorwarth, 2019, "Screening instruments for monitoring market power — The Return on Withholding Capacity Index (RWC).", Energy Economics 81: 227-237. doi:10.1086/599247.

[9] Mountain and Percy, 2019, *Market power in the National Electricity Market following the closure of the Hazelwood Power Station*, presentation to the Economics Society of Australia Annual Conference of Economists, July 2019.

3.3 Quantity-weighted offer price (action)

Each generator in the NEM submits an offer stack of up to ten price-quantity pairs. The quantity-weighted offer price technique simplifies the offer stack by collapsing it into a single value, which is averaged across the offered price tranches weighted by the quantity offered at each tranche.

The QWOP metric is used to track participant actions. It can be used to assess periods where generators have submitted higher bids. The New Zealand Electricity Authority has used QWOP in its market monitoring report to assess whether participant bids reflect costs.[10]

3.4 Combining the metrics

While the underlying metrics have been applied in other contexts, we extended these metrics by combining them.

3.4.1 Surplus capacity screen and QWOP increase (screen + action)

The Alberta MSA uses the surplus capacity metric to screen for potential instances of withholding before conducting a manual analysis of bids in outlier periods.[11] As we assess economic withholding in 4 NEM regions across multiple years, for practicality we performed further screening before conducting any manual analysis. To this end, we combined the QWOP increase analysis with the surplus capacity screen to assess whether participants systematically increase bids during outlier periods.

3.4.2 RWC and QWOP increase (incentive + action)

The New Zealand Electricity Authority has previously performed analysis comparing the QWOP to input prices to test whether offer prices deviate from costs.[12] However, performing this analysis without access to direct cost information requires significant cost modelling. In lieu of this, we measured the correlation between the QWOP and the estimated incentive to withhold (RWC). Our objective was to assess whether specific participants or generators consistently increase bids (as measured by QWOP) during periods when we estimated they have a stronger incentive to do so. If they do, this presents suggestive evidence of potential economic withholding.

[10] Electricity Authority (NZ), *Market Monitoring Review of Structure, Conduct and Performance In The Wholesale Electricity Market Since The Pohokura Outage In 2018,* October 2021.

[11] Alberta Market Surveillance Administrator, *Supply Cushion Methodology and Detection of Events of Interest, Directions Paper*, June 2012

[12] Electricity Authority (NZ), *Market Monitoring Review of Structure, Conduct and Performance In The Wholesale Electricity Market Since The Pohokura Outage In 2018,* October 2021

4 Surplus Capacity

4.1 Surplus Capacity Overview

Starting with definitions:

- **Surplus capacity:** the non-dispatched capacity in a market that is available to generate. This is the total available generation minus the current load.
- **Surplus capacity-price *relationship*:** the relationship between the dispatch price and surplus capacity. We expect the dispatch price to be typically higher when the surplus capacity is lower (as more expensive generating sources are required), though this relationship is not typically linear.
- **Surplus capacity *screen*:** a method to screen for periods when the dispatch price significantly exceeds the typical price given the surplus capacity.

Mathematically, the domestic surplus capacity, for state *s* in period *t*, SC_{st}, is:

$$SC_{st} = C_{st} - L_{st} \tag{1}$$

Where for a period t and state s, C_{st} is the total *effective* generation capacity[13] and L_{st} is the load.

Figures 6.1, 6.2 and 6.3 in the performance report were created using the methodology set out in this section. Note that to create Figures 6.1 and 6.3, we did not split the sample by wind and solar generation share.

4.2 Methodology for surplus capacity

To implement the surplus capacity-price relationship and screen, we undertook the following steps:

- split sample by wind and solar generation to account for how these types of generation are changing the relationship between surplus capacity and price (section 4.2.1)
- defined rolling bins to account for non-linearities in the merit order (section 4.2.2)
- trimmed sample to reduce effect of outliers on the screen (section 4.2.3)
- screened for outlier periods using the mean + 3 standard deviations for each rolling bin (section 4.2.4)
- defined a domestic surplus capacity (section 4.2.5).

4.2.1 Split sample by wind and solar generation

Greater intermittent generation by low-cost wind and solar generators means that more capacity is submitted at low prices. This changes the structure of the merit order and therefore surplus capacity-price relationship. To account for this dynamic, we split each

[13] Capacity that is available and unconstrained. Our process for calculating this is defined in Appendix D.

market (defined as all dispatch periods within a financial year for a NEM region) by the level of wind and solar generation.

To do this, we collected all dispatch periods within a NEM region and grouped those dispatch intervals into three categories based on the share of intermittent generation.[14] Appendix A presents the percentage share of dispatch periods by category of intermittent generation for each region and year. The categories are:

- low (below 5%)
- medium (5% to 25%)
- high (above 25%).

We estimated a separate surplus capacity-price relationship for each category in each region for each financial year. For example, for NSW, we separately estimated a relationship for periods of low, medium and high intermittent generation in 2021–22. Note that we only estimated a separate relationship if there were at least 4,000 observations: if there were insufficient observations to define a relationship, we allocated the additional observations to the next category.[15]

Our approach relative to alternatives/extensions: For simplicity, we only used three categories. The choice balanced being able to identify differences in the relationship with having enough data to sensibly plot the supply cushion. At cost of added complexity, it would be possible to extend our approach by specifying the surplus capacity as a regression equation to allow for complicated relationships between variables such as renewables share, seasonality and time of day.

4.2.2 Define 'rolling bins' to account for non-linearities

For each observation, we defined a 'rolling bin' that contains the 5,000 neighbouring observations in terms of surplus capacity level. For a given observation, the rolling bin combined the nearest 2,500 observations with a higher surplus capacity and the nearest 2,500 observations with a lower surplus capacity. For each rolling bin, we then calculated the mean (and standard deviation) of the dispatch price. Using a rolling bin meant we could recover a non-linear relationship between the dispatch price and surplus capacity. To visualise, we refer to Figure 5, which plots the surplus capacity-price screen in NSW for each financial year at different levels of wind and solar generation.

We judged that the choice of 5,000 observations led to an appropriate level of smoothness. Defining larger bins seemed to create over-smoothing, causing us to miss important changes in the relationship by averaging over too much data. Defining smaller bins led to kinks as we

---

[14] A market typically has 105,120 observations, which equates to 12 five-minute dispatch periods per hour; 24 hours day; and 365 days in a year

[15] To be specific, if there are fewer than 4,000 observations for the category of high intermittent generation, then these observations enter the medium category (e.g., NSW for 2017-2018). If there are fewer than 4,000 observations for the category of medium intermittent generation, then these observations enter the low category (e.g., QLD for 2017-2018). If there are fewer than 4,000 observations for the category of low intermittent generation, then these observations enter the medium category (e.g., SA for 2022-2023).

did not average over enough data and therefore introduced too much noise into the mean and standard deviation of price for each bin.

Our approach relative to alternatives/extensions: We used rolling bins defined by the number of neighbouring observations, which allowed us to capture non-linear relationships (for example, a more sensitive surplus capacity-price relationship at the point where the generation mix typically changes). A natural alternative would be to define fixed bins of uniform size – for example, using bins of 500 MW.[16] However, using bins of fixed width leads to discrete jumps in the screen from one bin to another. Additionally, it would require significant manual work to select bins of the correct size and curate the data to arrive at sensible results. By contrast, the use of rolling bins allowed us to automate this step, which enabled us to perform the analysis across multiple regions and years.

4.2.3 Trim sample to reduce effect of outliers on the screen

We trimmed the highest 0.01% of price observations from each rolling bin before calculating the mean and standard deviation. The purpose of trimming the data was to reduce the impact of statistical outliers, which might affect the surplus capacity-price relationship.[17] Without trimming, high price observations (such as prices at the cap) would increase the standard deviation and might cause the screen to miss some genuine outliers.

Our approach relative to alternatives/extensions: Alternative approaches would be to not trim the sample for outliers at all (in which case the screen would be higher and would identify fewer outlier price periods), or to trim more of the sample (in which case the screen would be lower and would identify more outlier periods). We judged that trimming 0.01% of outlier periods struck an appropriate balance.

4.2.4 Screen using mean + 3 standard deviations

Our surplus capacity screen is intended to detect price outliers, which are periods when the dispatch price is significantly higher than expected given the surplus capacity-price relationship. To do so, we applied a screen, which for each rolling bin, is the mean dispatch price plus 3 standard deviations (after trimming for outliers). The choice of 3 standard deviations follows prior precedent from Alberta MSA.[18] For examples of our outlier screen, we refer to Figure 3, Figure 5, Figure 7 and Figure 9 in Appendix B.

Our approach relative to alternatives/extensions: our approach of using 3 standard deviations follows a standard used by regulators overseas. Using a higher standard deviation would identify fewer observations, while a lower standard deviation would identify more.

4.2.5 Interconnector flows

Our analysis was performed at a region level and used the domestic surplus capacity-price relationship. This meant we only considered domestic generation and load for each region

[16] This approach would require us to find the mean and standard deviation price for all periods with a surplus capacity ranging from 0 to 500 MW, from 500 to 1000 MW, and so on. For an example of using bins of fixed width, see Alberta MSA, *Supply Cushion Methodology and Detection of Events of Interest, Directions Paper,* June 2012.

[17] Note that we only trim the sample for the purpose of determining the surplus capacity-price relationship and the surplus capacity screen – we of course include all outlier observations when we perform subsequent analysis.

[18] Alberta MSA, *Supply Cushion Methodology and Detection of Events of Interest, Directions Paper,* June 2012.

and did not consider the effect of interconnector flows. In cases where domestic load exceeded domestic generation (which might occur if the region is importing generation to cover a shortfall) we set the surplus capacity to zero.

Our approach relative to alternatives/extensions: Alternative methods involve adjusting the domestic surplus-capacity relationship to incorporate interconnector flows in the following ways:

- **Actual flows:** add/subtract the actual generation that flows over the interconnector
- **Potential flows:** use interconnector limits and flows data to add generation that could potentially be drawn upon to flow over the interconnector (or to subtract flows that are forced out of a region).[19]

We extend the equation for the domestic surplus capacity (see equation (1)) to account for actual and potential interconnector flows:

$$SC_s = \underbrace{(C_s - L_s)}_{Domestic} + \underbrace{\left(N_{s' \to s} + \tilde{N}_{s' \to s}\right)}_{Foreign} \quad (2)$$

Where for dispatch period *t* and state s, C_s is the total domestic capacity; L_s is the domestic load; $N_{s' \to s}$ is the interconnector flow into state *s* from foreign state s' (+ is inflow, - is outflow); and $\tilde{N}_{s' \to s}$) is the potential flow that state s could draw from state s'.

Figure 10, Figure 11, Figure 12 and Figure 13 compare the surplus capacity-price relationship we observe using each of the three interconnector treatments (no interconnector flows, actual flows and potential flows). The results show that accounting for actual and potential interconnector flows typically does not lead to a material change in the relationship between surplus capacity and price. Given the relationships are similar across the three possible treatments, we chose the simplest option. Further analysis would be needed to determine the value of incorporating these additional features into our analysis.[20]

4.3 Data requirements for surplus capacity

To calculate the level of surplus capacity, we required the following data for each dispatch period and NEM region:[21]

- native demand (MW)
- total effective capacity (MW)

In addition, to calculate the surplus capacity-price relationship and the surplus capacity screen, we required dispatch price data ($ per MW). Moreover, the process for calculating total effective capacity required its own data inputs and is outlined in Appendix D.

---

[19] This is the approach we use for our pivotal supplier test as laid out in AER, *Wholesale electricity market performance report 2022–Methods and assumptions*, December 2022.

[20] For example, assessing the extent to which changes in the dynamic limits that determine the potential interconnector flows actually reflect underlying changes in the merit order and therefore affect the dispatch price.

[21] Noting that accounting for interconnector flows would require data on these flows.

Since we have a surplus capacity and a dispatch price observation for every 5-minute period, the supply cushion-price relationship in each market was built from 105,120 observations. This breaks down into 12 five-minute dispatch periods per hour; 24 hours in a day; and 365 days in a year.

5 Returns from withholding capacity (RWC)

5.1 RWC overview

The returns from withholding capacity (RWC) measure estimates the extent that a participant increases its profit from withholding capacity. To be specific, this profit is an estimate of the increase in the dispatch price from withholding multiplied by the size of the participant's portfolio that is currently dispatched (and so stands to benefit from the higher price). Mathematically, the RWC of participant *i* in period *t* is the product of a price and portfolio effect:[22]

$$RWC_{it} = \underbrace{\Delta p_t}_{Price\ effect} \cdot \underbrace{(Q_{it} - 10)}_{Portfolio\ effect} \quad (3)$$

Where:

- $\underbrace{\boldsymbol{\Delta p_t}}$ **= price effect** is the extent that withholding 10 MW increases the dispatch price in period *t*.[23]
- $\underbrace{\boldsymbol{(Q_{it} - 10)}}$**= portfolio effect** is the generation portfolio that benefits from the estimated price increase from withholding 10MW. We consider both a 'raw' portfolio, which is the total in-merit capacity of participant *i* in period *t*, and additionally a contract-adjusted portfolio.

We used the methodology set out in this section to develop Figure 6.4 in the performance report.

5.2 RWC methodology

5.2.1 Price effect

We estimated the price effect using the slope of the surplus capacity-price relationship. This allowed us to estimate the typical increase in the dispatch price due to a reduction in capacity offered (or an increase in load). We estimated this relationship separately for low, medium and high levels of wind and solar output in each market. In each case, we used the relationship of mean dispatch price per rolling bin (as explained in section 4.2.2). Next, we fitted a smooth curve to this relationship, so that the slope was the derivative of this curve.[24] This approach maintained consistency between the surplus capacity screen and the estimated price effect for the RWC incentive. As a result of this process, we estimated a

---

[22] Note: the calculation assumes that the marginal cost of the withheld capacity is equal to the dispatch price. The implication is that the lost profit of the withheld capacity is zero. If the marginal cost of the withheld capacity is less than the dispatch price, then the RWC will be lower. While this does affect the interpretation of the *level* of the RWC, it should have less effect on *changes* of the RWC (that is, whether the incentive to withhold has increased or decreased).

[23] Note: we divide the $ per MWh price increase by 12 to calculate the profit increase for 5 minutes (rather than for one hour). Note also there is no *i* subscript for the price effect as all participants benefit equally from a price increase.

[24] Specifically, we apply a symmetric nearest neighbor linear smoother using 500 observations. In effect this smooths the relationship between the mean dispatch price and surplus capacity to remove noise. In stata software, this is the "running" command. Other approaches would lead to similar results.

price effect for every dispatch interval. To be specific, there is a distinct price effect that depends on:

- the surplus capacity-price relationship for a region, financial year and category of wind and solar generation
- the surplus capacity in that dispatch period

where combining the two provides the slope and so the expected price increase from removing capacity from the merit order.

Our approach relative to alternatives/extensions: Our approach of estimating the price effect directly from the surplus capacity-price relationship maintained consistency between the surplus capacity screen and the estimated price effect for the RWC incentive.[25] In effect, our approach estimated the *average* increase in price from a reduction in capacity offered (or an increase in load). This averaging approach was designed to be pragmatic and capture general trends. Two extensions would increase complexity but could improve accuracy:

- **More detailed surplus capacity-price relationship**. To be pragmatic, our surplus capacity metric only differed by financial year and by wind and solar generation share. It would be possible to extend this by calculating additional surplus capacity-price relationships based on:[26]
  - more granular categories of wind and solar share (that is, more than 3)
  - time of day (for example, peak versus non-peak)
  - seasonal changes (for example, quarterly or monthly).
- **Using the bid stack.** Our approach recovered an average relationship between surplus capacity and price. This average typically differs from the exact incentive within a period. Therefore, one extension would be to use the actual bid stack to estimate the period-specific (rather than average) price effect. There are costs to this approach:
  - It is computationally challenging, because estimating the incentive for each period requires using bid-stack data in each period to then calculate how different bids would affect the market clearing dispatch price.
  - It would require accounting for disorderly bidding. In particular, the incentive for a specific period may depend on interregional and/or settlement constraints.[27] This would add additional modelling complexity.

5.2.2 Portfolio effect

Our analysis calculated the portfolio effect in two ways:

[25] Alternatively, at the cost of additional complexity, it would be possible to estimate the price effect by regressing the dispatch price on the surplus capacity. If adopting this approach, it would make sense to follow a similar approach for the surplus capacity-price relationship.

[26] It would be possible to do this following our approach or by specifying a regression type approach as set out for the alternatives.

[27] Prior to 5-minute settlement, calculating the incentive in a dispatch period is complicated by the fact that settlement occurs over a 30 minute period. Therefore a plant needs to consider the likely prices across all dispatch periods within the 30 minute period.

- **Raw**: Total capacity dispatched during the period. This assumes that all of a participant's in-merit capacity would benefit from a price increase. However, we know that in practice participants contract a significant amount of generation.
- **Contract adjusted**: Adjusts the capacity dispatched to account for the capacity that is plausibly contracted. This approach assumes that only capacity above the contracted capacity benefits from a price increase.[28] While such an adjustment is likely to better reflect actual incentives, contracting data is confidential which requires us to estimate the proportion of a participant's portfolio that might plausibly be contracted.

To adjust for plausible contracting, we inferred that a participant has some level of contracted capacity if it consistently dispatched above a certain capacity. As an example, if a participant consistently dispatches at least 1,000 MW at a particular time-of-day then it is plausible to consider this capacity is contracted.

For this consistent dispatch metric to be useful, the running capacity should be relatively stable. This is likely to be the case for a participant with significant capacity generated from baseload generators (such as coal). Conversely, a participant that dispatches less consistently is less likely to be contracted for a specific quantity of capacity. This is more likely to be the case for a participant with fewer baseload generators (for example, a participant with primarily hydro generators). Note that we applied the same approach for all participants, but our approach is likely to more accurate for participants with more consistent dispatch.

For this analysis, we assumed that a participant contracts on a monthly basis and that the contract is for delivery at a specific hour. For instance, a participant is contracted to deliver a specific quantity of generation at 6pm for the month of January.[29] For a given participant and for each month and hour of day there are roughly 360 observations (12 periods per hour multiplied by 30 days per month). We sorted the hour-month observations to calculate the 5th percentile (meaning that dispatch is higher for 95% of periods and lower for 5% of periods). We focussed on the lower percentiles because we presumed a participant would attempt to dispatch its contracted capacity most of the time and sometimes dispatch more when prices are sufficiently high.

---

[28] A participant's incentive to withhold likely depends on the extent that it has pre-sold in the forward market ("contracted"). To see this, consider 3 different levels of contracting:

- **Uncontracted:** the plant fully benefits from any increase in the spot price – that is, effectively the "raw" RWC
- **Half contracted:** first 50% of capacity sold at a fixed price. Therefore, a price increase only benefits the plant for any generation above this 50%
- **Fully contracted:** the plant receives zero benefit from a higher spot price.

Contracting affects incentives in 2 ways. First, a plant has a greater incentive to withhold its uncontracted capacity (as only this capacity benefits from the price increase). Second, a plant has an incentive to bid competitively for its contracted capacity because it has an obligation to fulfil (as otherwise it needs to purchase this generation from the spot market). Note that in the long run, even a contracted participant may have an incentive to withhold if sustained increases in the spot price allow it to negotiate higher priced forward contracts.

[29] We understand this does not precisely align with contract definitions in the NEM (for example, ASX has monthly baseload contracts but only quarterly peak contracts). Our approach should be robust to longer contracts (e.g., quarterly) or if contracts are not by the hour. This is because we are estimating at a more disaggregate level. If contracts were less disaggregated, then we should observe little change between hours of day or between months.

Alternatives/Extensions: Our approach inferred plausible contracting using consistent dispatch. A preferable approach would be to use actual contracting data if it were available. In the first instance, access to actual contracting data would allow us to directly adjust the portfolio effect. In the second instance, having this data could inform whether our consistent dispatch approach (or any other approach to infer contracting) is accurate.

5.3 Data requirements for RWC

Estimating the price effect of the RWC required the same data as used for the surplus capacity-price relationship. That is, for each period and state we required:

- dispatch price ($/MWh)
- native demand (MW)
- effective capacity (MW).[30]

Estimating the portfolio effect (both raw and contract adjusted) required data on the quantity of generation dispatched per participant for each period. There is an RWC for each participant per 5-minute dispatch period. This amounts to 105,120 yearly observations of RWC per participant. This breaks down into 12 five-minute dispatch periods per hour; 24 hours in a day; and 365 days in a year. For simplicity, we focussed our analysis on the 5 largest participants in each region as defined by total generation output. Larger firms typically have greater incentive and ability to withhold capacity to influence the price.

[30] See Appendix D.

6 Increase in quantity weighted offer price (QWOP)

6.1 QWOP increase overview

Generators offer capacity into the market by submitting an offer stack of up to ten price-quantity pairs. The quantity weighted offer price (QWOP) simplifies a generator's offer stack into a single number. The QWOP is the average of the offer price, weighted by the quantity offered in each price tranche. The QWOP of a generator *i* in dispatch period *t* is:

$$QWOP_{it} = \sum_{J} \frac{q_{ijt}}{Q_{it}} \cdot p_{ijt} \tag{4}$$

Where *j* represents a price tranche. For a dispatch period *t*: q_{ijt} and p_{ijt} is the capacity offered and offer price per tranche and Q_{it} is the total offered capacity by the generator. By way of an example, a generator that offered 100MW at $20/MWh and 50MW at $50/MWh would have a QWOP of $30/MWh, which follows from $20*(100/150) + $50*(50/150).

Our analysis assessed how each generator increases its QWOP relative to its average QWOP in that quarter[31] and for that level of running capacity. Therefore, for each generator we initially calculated the QWOP per dispatch period. Then, we estimated the generator's average or typical QWOP. The generator's QWOP increase is the difference between the QWOP for the dispatch period and the average QWOP:

$$(QWOP\ Increase)_{it} = QWOP_{it} - \overline{QWOP}_i, \tag{5}$$

Where $\overline{QWOP}_i$ is a generator's normal QWOP. A positive QWOP increase value means the generator increased QWOP relative to normal while a negative value means it decreased relative to normal.

Alternatives/Extensions: Our approach was to calculate QWOP over the entire offer stack. We did this for simplicity as well as because it uses all tranches in the offer stack (each of which may be relevant to assess withholding behaviour). This approach, however, means that the QWOP may be affected by changes to bids that are unrelated to withholding and do not affect the dispatch price. For example, QWOP would fall if a generator were to bid some capacity at the price floor to avoid turning off. Equally, QWOP would increase if a generator were to bid some capacity at the price cap to avoid turning on. These changes may not necessarily affect the dispatch price. Two possible extensions could focus on:

- **QWOP for a window near the dispatch price**. For example, considering QWOP for offers between $0 to $500 per MWh for periods in which the dispatch price is between also within that range. This would focus analysis on QWOP increases that might slightly increase the dispatch price. However, this approach would require careful consideration of the correct window for both the QWOP and dispatch price. It would have the limitation of failing to capture QWOP increases outside the range of the window. Further, it would not capture instances when a generator significantly increases it bids (for example, to the price cap).

[31] Or to that month, in 2022.

- **QWOP for tranches above the dispatch price.** For example, if the dispatch price is $100 per MWh, this QWOP metric would focus on offers above $100 per MWh. While this approach would focus the analysis on bids that can theoretically increase the price, it would be more challenging to estimate an average QWOP given we would need to consider a typical QWOP at any given price.

6.2 Methodology

We calculated the average QWOP per:

- tranche of running capacity
- quarter (before 2022) and month (after 2022).

Alternatives/extensions: Calculating the QWOP increase relative to an average QWOP was a pragmatic choice that accounted for the two most relevant variables of seasonality and running capacity. A more complicated option would be a regression equation of QWOP, where a QWOP increase is the regression residual.

6.2.1 QWOP per tranche of running capacity

A generator faces costs as a result of ramping up and ramping down. Therefore, a generator typically offers different bids depending on its running capacity. For example:

- **Low running capacity:** submit some capacity at a relatively high price, to only turn on when economical
- **High running capacity:** submit some capacity at a relatively low price, to only turn off when economical.

Our approach was to estimate the average QWOP per tranche of operating capacity. To do so, we calculated the percent of a generator's maximum capacity that was operating in each period. Using this percentage, we defined 10 tranches of current operating capacity (that is, from 0% to 10%, 10% to 20% and so on to 90% to 100%).

6.2.2 QWOP per quarter and month

Operating costs vary over time because of fuel costs (including concerns around fuel scarcity) and operating conditions (for example, ramp up costs may change if market conditions change so generators are likely to remain on for more or less time). Therefore, we expect for generators to bid differently bid depending on the time of year. Our approach was to estimate normal QWOP by:

- **Quarter**: from 2017–2021, which strikes a good balance between capturing seasonal effects and the number of observations for estimation
- **Month**: for 2022, motivated by rapidly increasing fuel costs, which meant that the seasonal effect was changing much more quickly than before 2022.

6.3 Data requirements for QWOP increase

Estimating the QWOP increase required the following data:

- offer stack by generator for each dispatch period
- maximum capacity by generator (MW).

For participants that submit offers in every dispatch period of the year, we calculated 105,120 QWOPs per year for each station. This breaks down into 12 five-minute dispatch periods per hour; 24 hours in a day; and 365 days in a year.

7 Surplus capacity + QWOP increase (screen + action)

This analysis assesses whether specific generators consistently submit higher bids during outlier periods. There are two steps to this:

- **Step 1**: The surplus capacity screen identifies outlier periods when the dispatch price is significantly higher than expected given the level of surplus capacity (see section 4 for how we identified outlier periods)
- **Step 2**: For these outlier periods, we calculated the mean QWOP increase for a generator each year (see section 6 for how we calculated QWOP increases after controlling for running capacity and quarter/month).

Our analysis was at both the generator and fuel type level.

- **Generator level**: A generator may be withholding if it consistently increases its QWOP during outlier periods. Such behaviour may indicate that the generator is causing the outlier price periods by submitting higher bids
- **Fuel type level**: We averaged across all generators for a given fuel type, weighting by generation capacity to place more emphasis on QWOP increases by larger generators (that is, we gave ten times more weight to a generator of 1,000 MW than to a generator of 100 MW). As a robustness check, we further considered the average QWOP increase after removing QWOP *decreases* of more than $500 per MWh. This step was to indicate the possible impact of disorderly bidding, which occurs when a generator submits very low bids to ensure dispatch.

We used the fuel type level results to create Figure 6.6 in the performance report.

Our approach relative to alternatives/extensions: Since our approach is intended to be an initial screen, we considered the average QWOP increase. Possible extensions could involve looking at the distribution of QWOP increases in outlier periods or to compare QWOP increases to likely changes in the dispatch price. While these other analyses may be informative, this metric is intended to be initial screen to then combine with other metrics, so there is limited value to more complicated analyses.

8 RWC + QWOP (incentive + action)

In this analysis, we assessed whether specific generators consistently submit higher bids when we estimated that there is a stronger incentive to withhold capacity. Our 2 metrics are:

- **RWC**: an estimate of the participant's profit increase from withholding capacity. We refer to section 5 for details on how the estimated incentive depends on:
  - the level of surplus capacity
  - the share of wind and solar generation
  - the participants running capacity (including consideration of the potential effect of contracting)
- **QWOP increase**: the extent to which a generator submits a higher bid (see section 6 for details on how we calculate QWOP increases after controlling for running capacity and quarter/month).

Combining the two metrics allowed us to determine whether generators may be economically withholding when the profits to do so are estimated to be higher. We implemented this analysis at both the generator and participant level:

- **Generator level**: We calculated the linear correlation between QWOP increase and RWC for each generator in each financial year. The correlation coefficient is a value between -1 (perfect negative correlation) and +1 (perfect positive correlation). A positive correlation implies the generator typically submits higher QWOP when the (estimated) profit gain is larger and vice versa.[32]
- **Fuel type level:** We aggregated these generator-level correlations to a fuel-level summary by taking the average across all generators. In calculating the average, we weighted the generator correlation by station capacity. Weighting by generator capacity places more emphasis on QWOP increases by larger generators.

We used the fuel type level results to create Figure 6.5 in the performance report.

Our approach relative to alternatives/extensions: Since all limitations to the QWOP and RWC metrics would add noise – which would reduce the estimated correlations – we chose the simplest approach of estimating a linear correlation. However, non-linear relationships could be uncovered by using a more complicated regression approach.

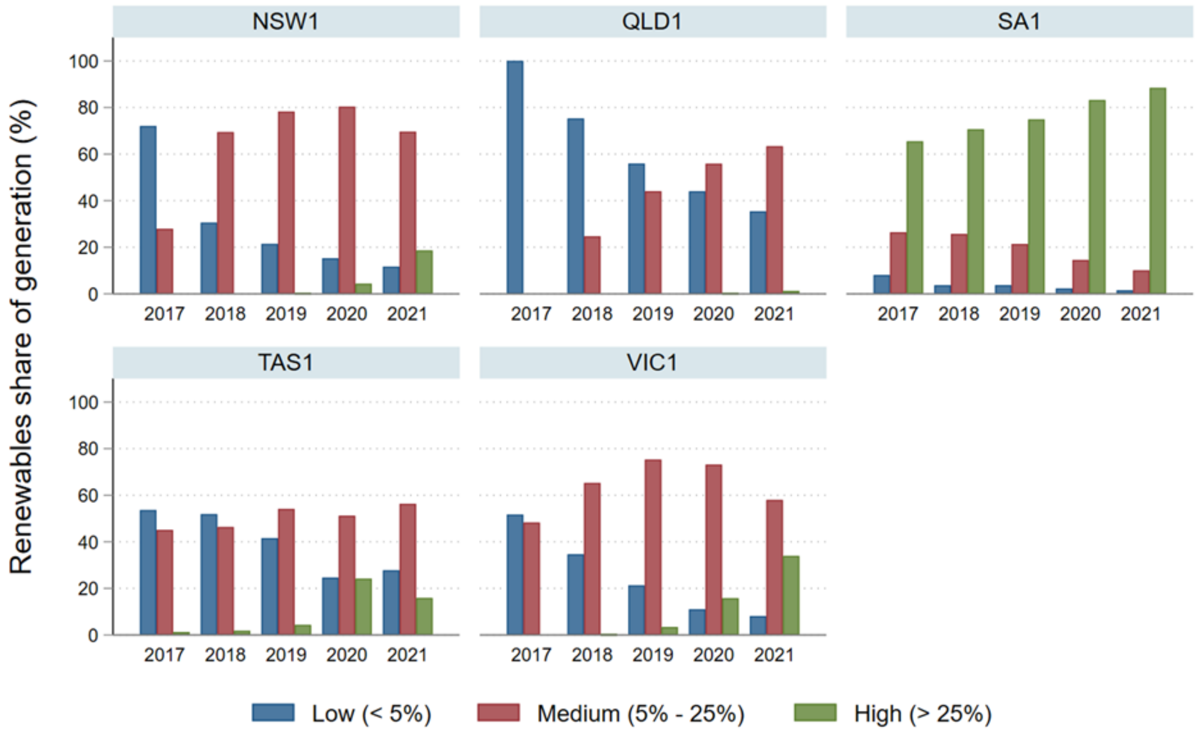
A further possible extension to the analyses in sections 7 and 8 is to combine them to apply the *incentive* + *action* assessment in outlier periods (*screen*). This analysis would assess whether generators consistently submit higher bids in those outlier periods when the incentive to withhold is greater. In other words, it would highlight the correlation between incentives and actions specifically for the periods when the dispatch price is known to be high. Nevertheless, implementing this analysis is likely to require an approach to measuring the price effect within each period rather than estimating the average price effect across a

[32] Due to the large number of observations (roughly 100,000 per station per year), we did not consider statistical significance. For example, any correlation larger in magnitude than 0.01 would be considered statistically significant at $p < 0.01$.

financial year. This is because our approach of averaging across all periods in a financial year is likely to provide a less accurate estimate of incentives when applied to outlier periods.

Appendix A. Share of wind and solar generation

Figure 1: Share of dispatch periods by category of wind and solar generation for each region and financial year



Note that these charts show financial year data: ‘2017’ represents 2017–18, ‘2018’ represents 2018–19, and so on.

Appendix B. Surplus capacity-price relationship and screen for each region

These charts show the surplus capacity-price relationship and screen for low, medium and high levels of wind and solar generation in each region. In some regions, there were insufficient observations (that is, fewer than 4,000) of a generation category in a year to define a surplus capacity-price relationship. In these cases, we have not plotted a relationship for that category of wind and solar generation share. Note also that the charts for 2022–23 are created from just one quarter of data (Q3 2022). See section 4.2 for a more detailed explanation of our approach.

Figure 2: Queensland surplus capacity-price relationship

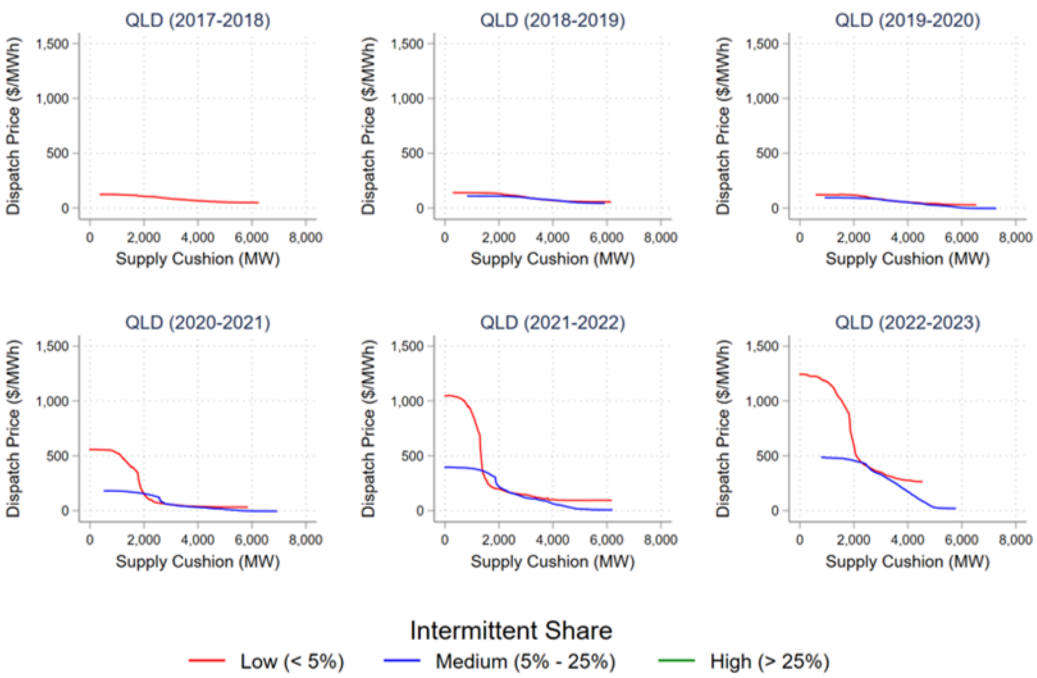


Figure 3: Queensland surplus capacity-price screen

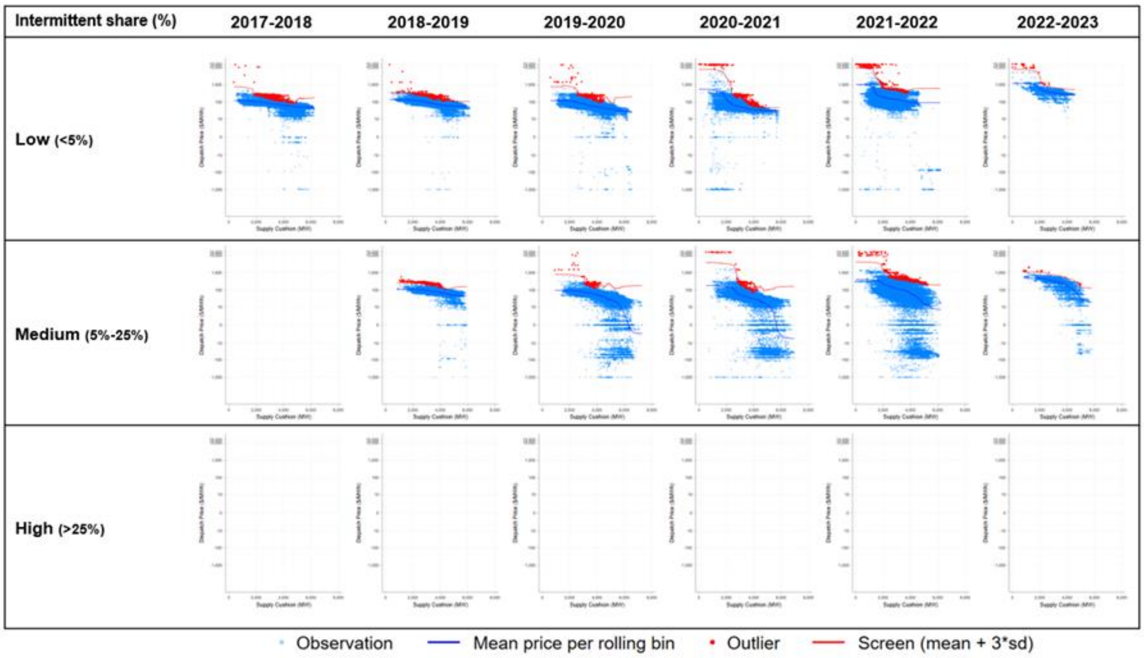


Figure 4: NSW surplus capacity-price relationship

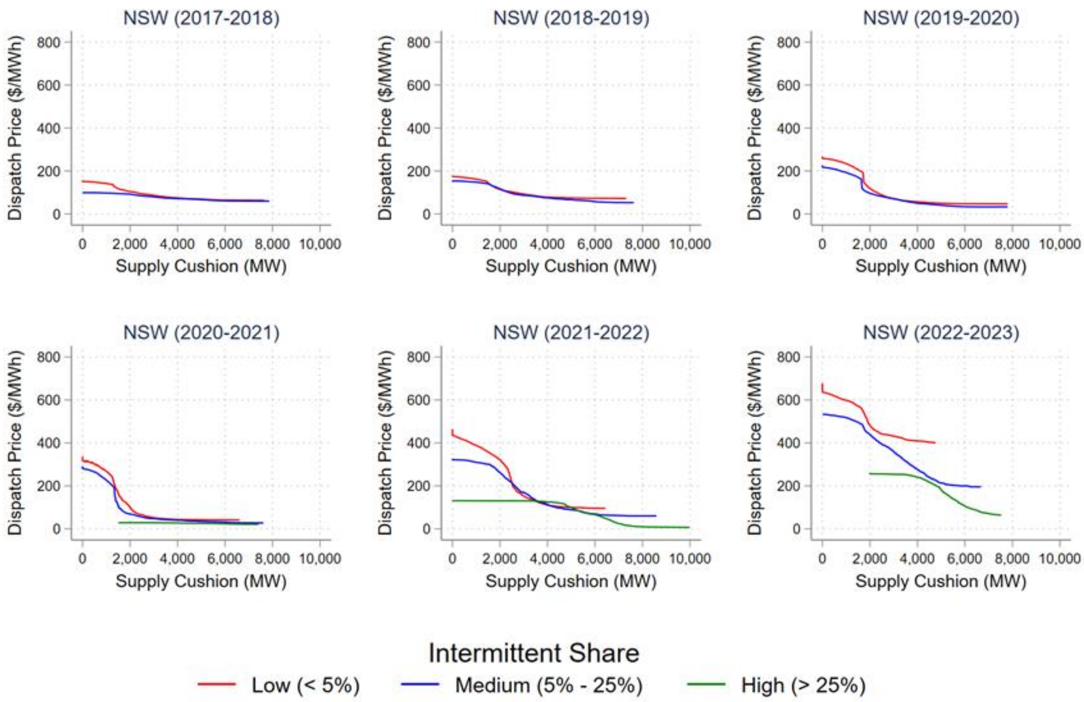


Figure 5: NSW surplus capacity-price screen

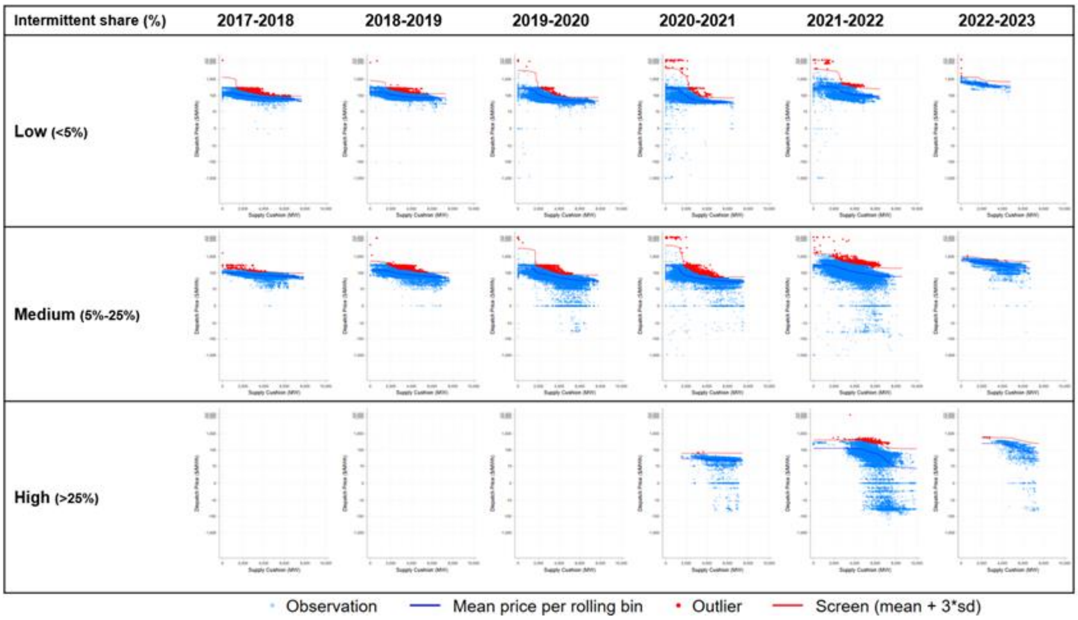


Figure 6: Victoria surplus-capacity price relationship

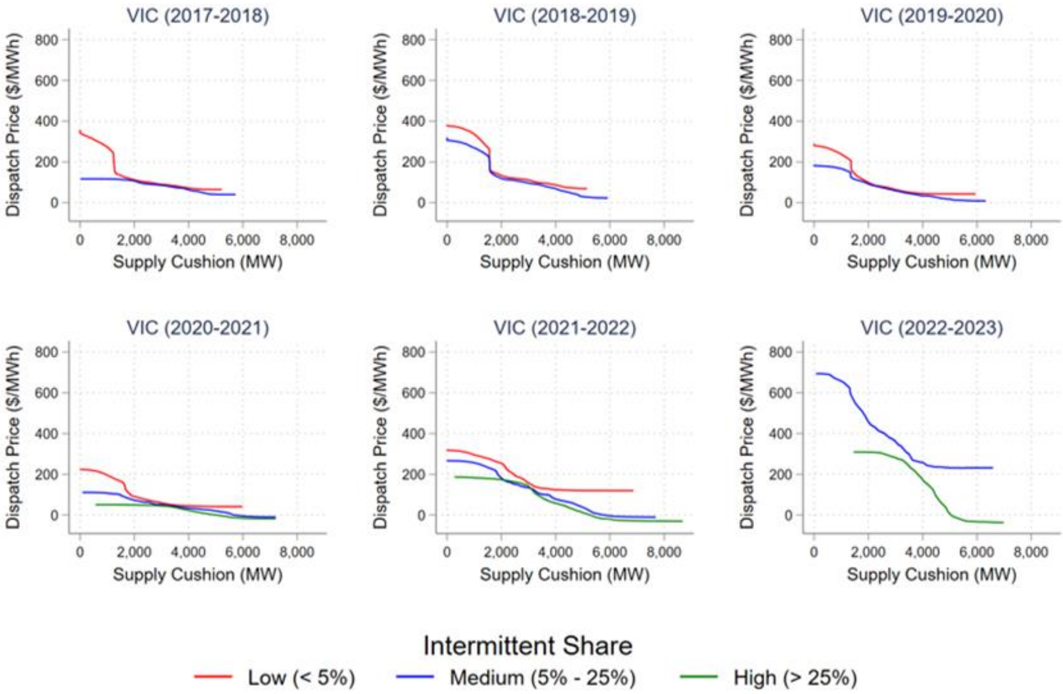


Figure 7: Victoria surplus capacity-price screen

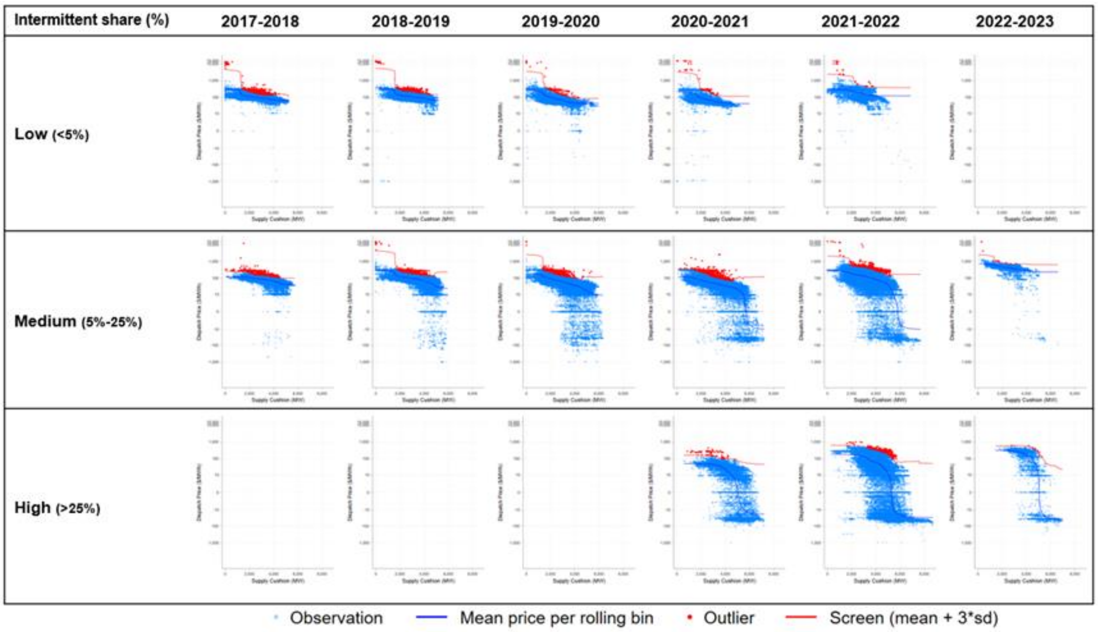


Figure 8: South Australia surplus capacity-price relationship

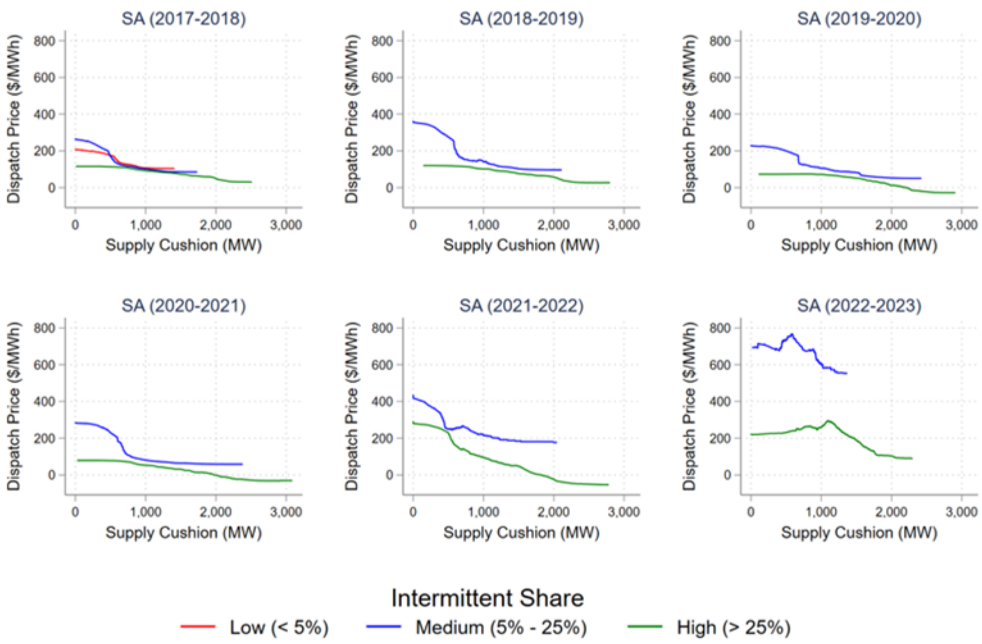
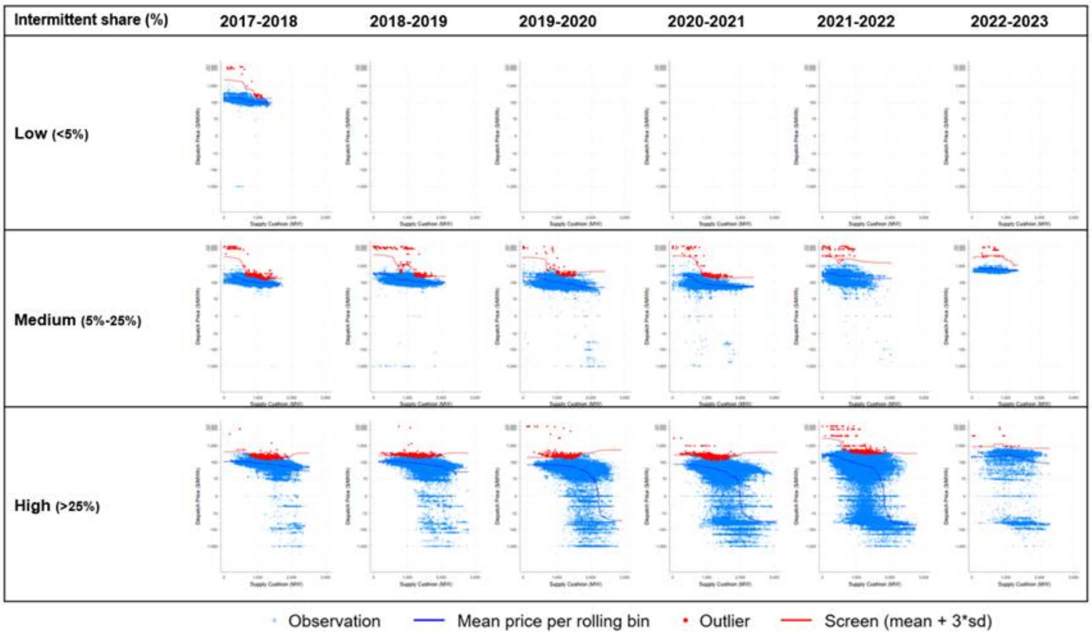


Figure 9: South Australia surplus capacity-price screen



Appendix C. Effect of different interconnector assumptions on the surplus capacity-price relationship

The following figures present the surplus capacity-price relationship for three options or assumptions for specifying the interconnector relationship (see Section 4.2.5 for further detail):

- **Option 1:** Domestic surplus capacity-price relationship
- **Option 2:** Adds actual interconnector flows
- **Option 3:** Adds actual and potential interconnector flows, as per the method used for the AER's pivotal supply test.

Figure 10: Effect of interconnector assumptions on the Queensland surplus capacity-price relationship in 2021–22

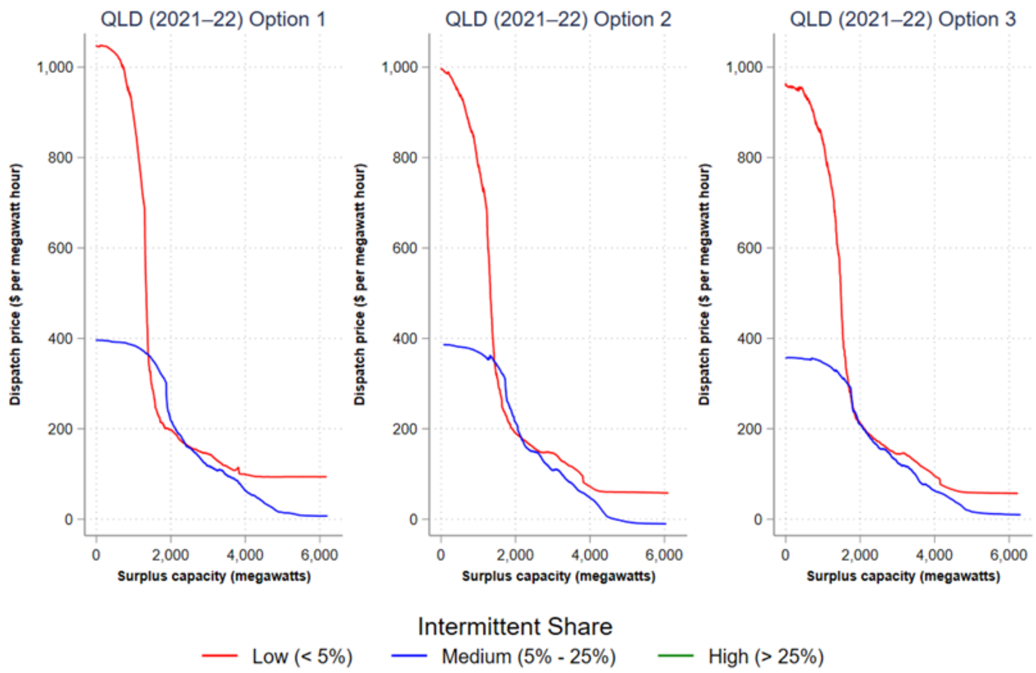


Figure 11: Effect of interconnector assumptions on the NSW surplus capacity-price relationship in 2021–22

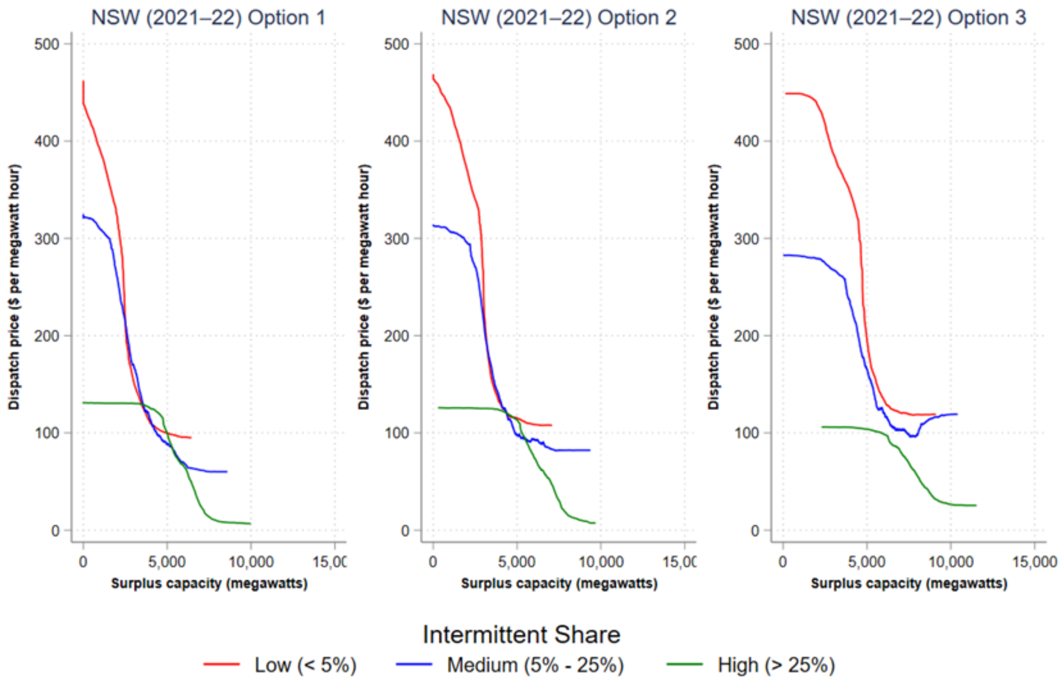


Figure 12: Effect of interconnector assumptions on the Victorian surplus capacity-price relationship in 2021–22

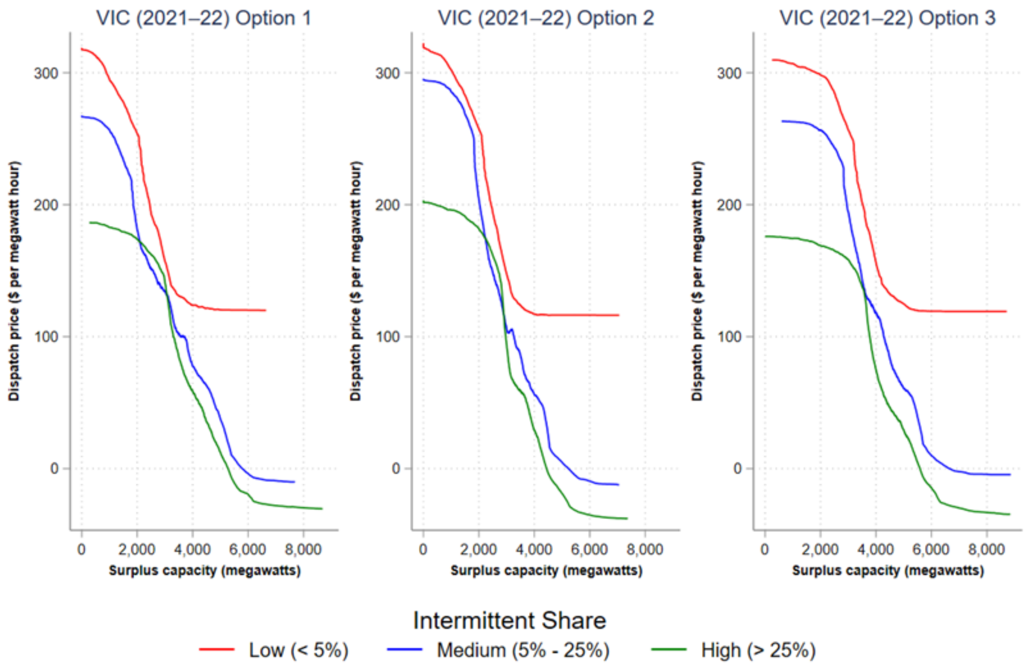
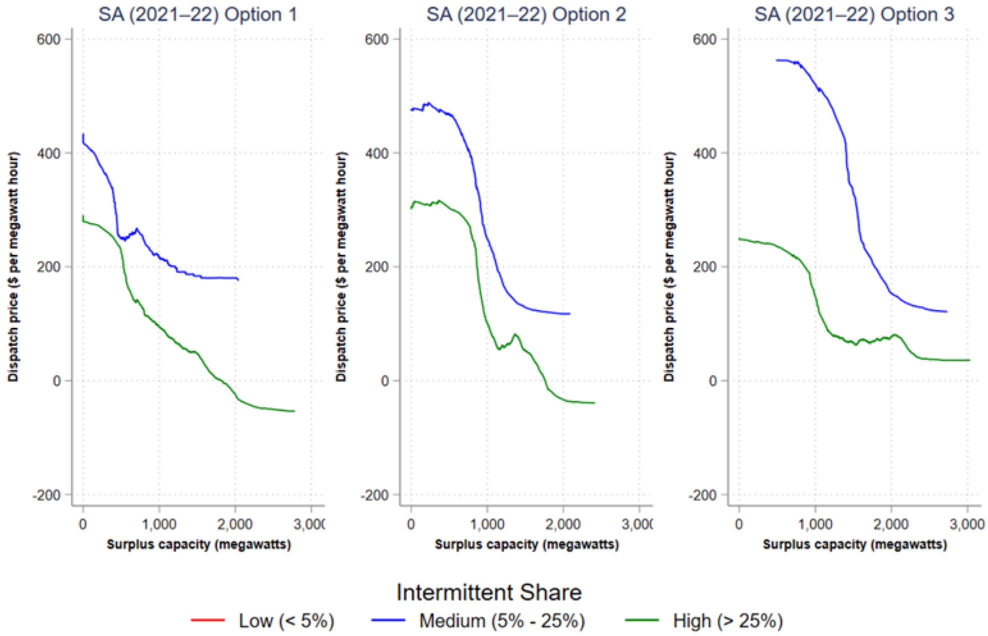


Figure 13: Effect of interconnector assumptions on the South Australian surplus capacity-price relationship in 2021–22



Appendix D. Calculating effective capacity

The effective capacity in a region is capacity that is available to generate and not constrained. To calculate this we sum the effective capacity of all scheduled and semi-scheduled generating units except for AEMO reserve trader units (which have DUIDs starting with ‘DG_’ or ‘RT_’). To calculate effective capacity for an individual unit, we require the following data:

- effective bid availability not constrained, for all 10 price bands
- fixed load
- max avail

If a unit has a fixed load in a period, then it has no effective capacity in that period. However, if it has no fixed load, we calculate the effective capacity for each unit by summing its effective capacity for each price band in the following way:

- If the price band divided by the loss factor for the unit is less than the dispatch price, and if the unit is not setting price in the region, then its effective capacity is the smallest of:
  - the total cleared target
  - max avail
  - the sum of bid band avail in that price band and all lower price bands
- But if the price band divided by the loss factor is greater than the dispatch price, or if the unit is setting price, then the effective bid availability for that price band is the smallest of:
  - max avail
  - the sum of band avail in that price band and all lower price bands.

Appendix E. Selected summary statistics

Table 3: Number of price outliers per region by financial year

	2017–18	2018–19	2019–20	2020–21	2021–22
NSW	1,219	1,441	1,192	1,638	2,536
QLD	1,104	1,500	850	960	1,958
VIC	1,443	1,612	1,093	1,217	1,466
SA	1,025	1,544	1,077	1,036	1,297

Note: Outlier prices were defined as being more than three standard deviations from the mean in price in the relevant rolling bin.

Table 4: Number of estimated correlations between QWOP increase and RWC

	Black coal		Brown coal		Hydro		Gas	
	Plants	**Total**	**Plants**	**Total**	**Plants**	**Total**	**Plants**	**Total**
NSW	5	25	0	0	5	21*	3	15
QLD	8	40	0	0	3	15	3	15
VIC	0	0	3	15	5	25	8	40
SA	0	0	0	0	0	0	8	37*
Total	13	65	3	15	13	61	22	107

Note: We only considered the 5 largest participants in each state, ranked by total generation. For each plant we estimate a yearly correlation, which typically led to 5 estimates.

*Plant closures meant there were fewer than 5 estimates

Table 5: Number of calculated QWOPs by region and financial year

		2017–18	2018–19	2019–20	2020–21	2021–22
NSW	Coal Black	522,199	525,600	515,899	522,350	523,642
	Hydro	545,220	502,872	475,653	472,885	553,155
	Gas	274,403	299,960	223,952	195,667	233,813
Queensland	Coal Black	803,522	830,456	771,404	790,107	795,373
	Hydro	294,547	303,864	310,933	301,049	278,457
	Gas	274,300	240,833	276,211	263,424	270,401

Victoria	Coal Brown	315,360	315,360	316,224	315,310	315,360
	Hydro	490,141	523,598	521,552	491,903	473,863
	Gas	763,845	746,025	738,460	732,678	737,124
South Australia	Gas	804,332	765,783	838,123	854,364	749,960