High-Impact, Low-Probability Events and the Framework for Reliability in the National Electricity Market

PREPARED FOR

The Australian Energy Market Commission

PREPARED BY

Toby Brown

Neil Lessem

Roger Lueken

Kathleen Spees

Cathy Wang

February 2019

THE Brattle GROUP

Notice

- This report was prepared for the Australian Energy Market Commission, in accordance with The Brattle Group's engagement terms, and is intended to be read and used as a whole and not in parts.
- The report reflects the analyses and opinions of the authors and does not necessarily reflect those of The Brattle Group's clients or other consultants.
- There are no third party beneficiaries with respect to this report, and The Brattle Group does not accept any liability to any third party in respect of the contents of this report or any actions taken or decisions made as a consequence of the information set forth herein.

Copyright © 2019 The Brattle Group, Ltd.

Table of contents

Executive summary

Background

The AEMC has asked us to examine the economic concepts of risk aversion and loss aversion, and to consider whether these concepts might be relevant for how the reliability framework in the National Electricity Market (NEM) addresses risks associated with high-impact, low probability (HILP) events.[1] These concepts were raised by AEMO in connection with a recent rule change request, where it stated that the expected unserved energy (USE) reliability standard in the NEM assumes risk neutrality, which is inconsistent with the concept of risk aversion.

Ordinarily, the price mechanism of the wholesale market (ie, the spot market) operates to maintain a balance between supply and demand. When electricity supplies are interrupted, this is almost always caused by technical faults on the networks. Occasionally, however, AEMO has to instruct load to be disconnected because there is insufficient reserves to maintain the balance between supply and demand. The latter type of interruption is caused by a lack of reserves rather than technical faults on the networks.

In the NEM, high wholesale market prices are intended to encourage both more supply to be offered to the system and consumers to reduce consumption during periods of peak demand. In this way, the price mechanism continuously balances supply and demand, without the need for load shedding—ie, achieving "wholesale-level reliability". It would however, be very expensive to build a system with so much spare capacity that load shedding would not be necessary under any circumstance. The "reliability standard" in the NEM recognizes this - the standard is that no more than 0.002% of electricity supply in each region in each financial year be expected to go unserved because of an imbalance between supply and demand. AEMO is able to contract for emergency reserves through the Reliability and Emergency Reserve Trader (RERT) mechanism if it anticipates that the reliability standard will not be met.

When forecasting wholesale-level reliability, AEMO estimates the *expected* amount of unserved energy (USE) in MWh by modelling a wide range of possible outcomes, each of which has an associated quantity of USE and a corresponding probability that it will occur. The expected amount of USE is the weighted average across the range of outcomes, where the weights are the probabilities.

---

1 We are not aware of a precise definition of a HILP event in the NEM. HILP events are those which, if they were to occur, would have a large impact in terms of the quantity of load lost (USE). For example, in AEMO's modeling of Victoria for 2018/19, out of more than 20 million hours modelled, fewer than 0.03% have interruptions. However, of these hours with interruptions, the average quantity of interruption was 363 MW, but 5% of the hours had more than 1 GW of interruption (see AEMO (2018), *The NEM Reliability Framework, Additional information from AEMO to support its Enhanced RERT rule change proposal*, November 2018, pp. 14–15.

In this report, we examine the concepts of risk aversion and loss aversion, and consider whether these concepts might imply that expected USE is not adequately capturing consumers' attitudes toward wholesale-level reliability. We focus specifically on whether it is appropriate to weight an estimate of the impact of an HILP event by an estimate of the probability that the event will occur, or whether some larger weight should be used to reflect risk aversion. Unconnected with risk and based on our observations in other jurisdictions, we also discuss the possibility that models of wholesale-level reliability may fail to capture adequately all of the relevant HILP wholesale-level reliability events. Some such events may be missing from the model, or the corresponding impact may be underestimated, or the corresponding probability may be underestimated. In addition, the risk of more serious outages connected with network faults may be elevated under conditions that give rise to load shedding to manage a wholesale-level reliability event.

Risk aversion

In many circumstances, individuals make decisions which are not consistent with maximizing expected financial gains (or minimizing expected financial costs). Rather, individuals seem to prefer certainty to uncertain/risky outcomes. For example, people often prefer to purchase insurance against risks such as accidental damage to their possessions, and electricity consumers often prefer a fixed-price retail contract to one that passes through volatile spot prices in the wholesale market. The expected value of an insurance policy (a possible future claim to offset an insured risk eventuating, multiplied by the probability of needing to claim, less the certain up front premium payment) is negative, since the insurance provider will want to recover the expected claims, plus a margin, when it sets the premium. In the second example, electricity consumers do not mind paying a premium above expected volatile spot power prices for a steady price over an extended period. People purchase insurance because they prefer the certainty of being insured, even though it costs money. This type of behaviour is termed "risk aversion".

Economic theory provides two different explanations for why individuals are risk averse. In relation to large risks, which could give rise to significant changes in the individual's total wealth, risk aversion is explained by the idea that individuals try to maximize expected utility (the satisfaction they get from money and the goods and services that money buys) rather than expected wealth. The utility of additional wealth declines as wealth gets larger. Expected utility theory is the traditional explanation for risk averse behavior, so much so that many economists use the terms interchangeably. When individuals choose to avoid smaller risks, this behaviour is better explained by the concept of "loss aversion", one of the tenets of behavioural economics. Loss aversion states that relative to their expectation for the future, individuals dislike the chance of a small loss more than they like the chance of an equally likely gain of the same magnitude. For example, individuals often prefer to buy insurance in circumstances where expected utility theory would predict that the risk is too small to be worth insuring, such as warranties on white goods.

Incorporating risk aversion into policy decisions – other jurisdiction examples

We have not found clear examples of policy decisions where an option was chosen explicitly taking risk aversion (or loss aversion) into account. In analysing options for managing climate change risks, risk aversion is sometimes addressed by reporting both expected (average) costs and benefits along with the costs and benefits associated with the tails of the risk distribution. This approach is often described as a way of reflecting concerns that society may be more interested in avoiding the worst potential outcomes of climate change, than just the average expected potential outcome – ie, that society may be risk averse. However, in most policy contexts, the standard approach seems to be to consider probability-weighted expected costs and benefits without giving special weight to risks associated with HILP events.

Wholesale-level reliability risk in the NEM

In the NEM the price in the wholesale market cannot be greater than the market price cap (MPC). If the MPC were equal to the (marginal) value of load lost during an interruption, and if customers at risk of interruption pay the wholesale price, then those customers would be indifferent between paying the MPC in a particular interval or having their load shed, irrespective of any risk aversion. However, this is not the case: almost all NEM customers are supplied under a fixed-price retail contract, and so are much better off if they are not interrupted (retailers conversely may be better off financially during an interruption, since they can avoid paying the MPC for electricity).

If consumers are risk averse in relation to wholesale-level reliability, then they might prefer to pay for an additional insurance or safety net mechanism to avoid interruptions, such as allowing for an emergency reserve (such as the RERT) to be procured even in circumstances where the current reliability standard is expected to be met. For example, while HILP events contribute little to an estimate of expected USE, due to their low probability, consumers might prefer that an emergency reserve mechanism be used to reduce the impacts of HILP if they were to occur, as a form of additional insurance. The current reliability framework does not provide such insurance because the RERT is only used when the current reliability standard is not expected to be met.

Impacts of HILP wholesale-level reliability events

When there are insufficient reserves in the NEM, the system operator will direct “load shedding”. A quantity of load will be disconnected, in order to keep supply and demand in balance. The number of consumers disconnected and the duration of the event will depend on the degree of the imbalance and also on which loads are disconnected (load is disconnected in “blocks” of different sizes). If the load shedding needs to continue for an extended period of time, the outage is “rotated” over different blocks of load. Thus individual consumers should not be off supply for extended periods, even if the load shedding itself does continue for a long period. Because of the rotation, the impact on any one customer should be relatively small. For example, the impact would be small compared to that of many network outages, which are can be longer lasting and/or more widespread geographically. In particular, even HILP wholesale-level reliability events are expected to be managed by rotating outages.

While rotating load shedding should mean that the impacts of HILP wholesale-level reliability events are relatively small (compared to other potential outage events), the risks of faults giving rise to extended and/or widespread outages may be elevated under load shed conditions. When load is being shed, the system is, by definition, in an unusual state and may be under stress.

Since these conditions are rare, the system operator and market participants have limited experience of operating under such conditions. The impacts on consumers of fault-related outages could be more likely to eventuate and much greater in magnitude if the system is already tight because such outages have the potential to cascade in ways that might (in an extreme case) affect a wide geographic area (possibly an entire state) and can last days rather than hours. Unless there are simultaneous network faults, or cascading outages caused by common stresses on the power system, a "system black" will not result. The impacts of a system black event are probably much greater, per MWh of USE, than a rotating load shed.

Reliability frameworks in other jurisdictions

We have reviewed the reliability frameworks in several US jurisdictions (PJM, ISO-NE, ERCOT) and Great Britain. None of them explicitly discuss risk aversion. However, all four jurisdictions either target a reliability standard that is much higher than one based on expected costs and benefits, or their mechanisms effectively procure more resources than needed to meet the standard, or both. For example, the current capacity demand curve in PJM's capacity mechanism is consistent with achieving a reliability level of about 0.01 loss of load events per year —ie, a 1% chance of an outage across the year (which is much more reliable than the 0.1 loss of load event standard). Analysis of the ERCOT system shows that the reserve margin corresponding to an "economic" level of reliability would be consistent with an outage risk of greater than 10% per year. ERCOT has a similar market design to the NEM (it is an energy-only market without a capacity mechanism), but ERCOT has an administratively-determined price adder in pricing intervals with elevated wholesale-level reliability risks. This adder is designed to deliver a slightly greater degree of wholesale-level reliability than justified on the basis of expected costs and benefits. In addition, ERCOT also has an emergency demand response reserve mechanism, and is able to sign temporary emergency reliability contracts with generators that would otherwise retire if they are needed to support system security.

While these jurisdictions do not explicitly discuss risk aversion in connection with their reliability frameworks, two other reasons for seeking more reliability are discussed. First, the modelling underpinning the reliability framework in these overseas jurisdictions probably does not adequately describe all of the possibly-relevant HILP wholesale-level reliability events. The model may not fully capture the impacts if the event occurs (including the incremental risk of a security event), or it may not accurately estimate the probability of the event, or some possible HILP events are not in the model at all. As a result, the system operator tends to procure additional resources. Second, system operators are often concerned about system security risks, and often there are resources that could be used to reduce both security risks and wholesale-level reliability risks. System operators may find it more convenient to use an existing mechanism designed to procure resources for reliability reasons to deliver resources that will be used to manage security risks.

An alternative explanation is that it is possible that reliability frameworks may reflect a type of principal–agent problem (principal–agent problems arise where one entity acts on behalf of another but incentives of the principal and agent are not perfectly aligned). In this case, institutions are acting on behalf of customers in relation to reliability, but these institutions may not always follow customer risk preferences. This could be due to a perception that insufficient reliability would be costly for both institutions and customers, whereas procuring additional resources to increase reliability imposes modest additional costs on customers (and no additional cost on the institutions).

Conclusions

In all four of the overseas jurisdictions that we reviewed, the reliability frameworks ultimately resulted in the system operator procuring more resources than system modelling shows is needed to meet the reliability standard. It is not clear why this is the case, but possible explanations include: system models may not capture the full range or extent of HILP events; system operators may be using these reliability resources to address system security risks; and/or system operators and policy-makers may have a bias towards delivering additional reliability, for example because these institutions do not themselves bear the costs of purchasing additional reserves. If HILP wholesale-level reliability risks are associated with elevated risks of wider system outages for network or security reasons, we would recommend that security needs be addressed directly, because the type of resource or service that can most efficiently address security needs may be different from that which can efficiently address wholesale-reliability needs. The impacts on consumers of network faults and security events can be much larger than the impacts of even HILP wholesale-reliability events, so measures which reduce the probability or impacts of network and security events are more valuable than measures which address wholesale-level reliability events alone.

If rotating outages operate as planned and there is no additional security risk, it seems unlikely that HILP wholesale-level reliability events would have large impacts on consumers. Consumer preferences over wholesale level reliability risks will depend on the magnitude of the potential impacts that they face. Since the impacts of HILP wholesale level reliability events are relatively small on a per customer basis, there is no need to account for wealth-based risk aversion in measuring expected USE. It is however possible that consumer preferences in relation to wholesale-level reliability risk might reflect loss aversion, another type of risk aversion, which is observed in other contexts in relation to small losses.

If consumer preferences and expectations about wholesale-level reliability in the NEM include loss aversion so that they would prefer to avoid incremental reliability risk, then insurance would be valued by those consumers. However, there is no obvious way for consumers to signal their preferences because (so far as we know) there are no insurance-type products that cover the risk of interruptions caused by wholesale-level reliability events. Similarly, the Australian Energy Regulator's forthcoming update of consumer VCRs, focuses on valuing lost load conditional on an event having occurred, but does not assess consumer attitudes towards risk. While some consumers may value reliability sufficiently highly that they install backup generation or a backup battery, such systems can cover both common network-related outages and much less common wholesale-level reliability outages. Thus, decisions to install such systems are not driven by preferences about wholesale-level reliability. We do not know whether consumers in the NEM are risk averse in relation to wholesale-level reliability. It might be possible to assess consumer preferences through surveys and directly asking about willingness to pay for insurance against wholesale-level reliability events, but we are not aware of any such surveys. We would expect the results to depend on how consumers perceive these risks and what "baseline" level of outage risk consumers would use to assess options.

If survey results indicate a material preference for additional insurance, then adjusting the reliability framework to deliver additional reserves is one form of insurance mechanism that could be implemented in order to address loss aversion. Another form of insurance mechanism would be to pass on to customers that are interrupted the avoided costs of the energy that was not supplied to them and which, in consequence, their supplier did not have to pay for (but would have paid for if the customers' load had not been shed).[2]

[2] A similar concept was discussed by the AEMC in *Wholesale Demand Response Mechanism, Consultation Paper*, November 2018, Appendix D.

I. Introduction

A. Background

Many consumers place a high value on a reliable supply of electricity. In this context, a reliable supply means that consumers can use electricity at any time, and supply interruptions are infrequent and, when they happen, are localised, and do not last long.[3] Most supply interruptions are caused by faults on the network.[4] However, some interruptions are due to demand for electricity at a particular point in time exceeding the total amount of generating capacity available at that same point in time.[5] Ordinarily, at the wholesale level the price mechanism operates to "clear" the market and to keep the total quantity of generation equal to the total quantity of demand in each dispatch interval. Under some circumstances, the system operator may be able to call on reserves of generation or demand response, outside the market mechanism. However, if the market cannot be cleared using the price mechanism, and these out-of-market reserves are not available or exhausted, a mis-match between supply and demand occurs, and the system must be held in balance by the system operator deliberately cutting off supplies to some consumers. This paper discusses reliability only in the context of the latter type of interruption, though the same principles can be applied to the impact of network faults on reliability. We use "wholesale level reliability" to refer to the ability of the market mechanism (and out-of-market reserves available to the system operator) to achieve a balance between available generation and total demand (taking into account network constraints and demand response).

In the National Electricity Market (NEM), the primary mechanism for delivering wholesale-level reliability is the market itself. In a pricing interval where there is relatively little excess of available generating capacity over the likely level of demand, the market clearing price will be very high. If there are many intervals with high prices, market participants will see a signal that generating capacity is valuable and will consider investing in additional capacity. Likewise, loads will see that peak consumption is expensive and will consider reducing consumption at peak times (ie, demand will respond to the high prices).

---

3 Formally, a "reliable" power system is defined to have "adequate amount of capacity (generation, demand response and network capacity) to meed consumer needs". See AEMC (2018), *Reliability Frameworks Review, Final Report*, July 2018, p. i.

4 Between 2007-08 and 2016-2017, 95.63% of supply interruptions were distribution network interruptions; 0.94% were transmission network interruptions. See AEMC (2018), *Reliability Frameworks Review, Final Report*, July 2018, Figure 2.1.

5 Generating capacity includes in-market reserves.

The NEM clearing price cannot exceed the market price cap or MPC (currently set at $14,500/MWh), and this is also the level at which prices are automatically set when there is insufficient capacity to clear the market, i.e. when there is involuntary load shedding. The level of the MPC is chosen such that a plant which runs only a few hours per year is still forecast to be able to cover all of its costs. However, once that plant is running, any further increases in demand (or further generator outages) will have to be managed through dispatching out–of-market reserves, and, ultimately, load shedding (ie, outages for firm loads). The level of the MPC is set to deliver the NEM "reliability standard", which is that interruptions due to wholesale-level reliability are expected to result in load shedding equivalent to no more than 0.002% of total energy demand in any NEM region in any year. When this standard is met, "unserved energy" (USE) is expected to be no more than 0.002% of total demand in future years.[6] Modelling work is done to show what level of MPC will support a system that meets the 0.002% expected USE reliability standard.[7] In its most recent review of the reliability standard and settings, the Reliability Panel said:[8]

> *The present market price cap and cumulative price threshold have been, and are likely to continue to be effective at limiting market participants' exposure to excessive high prices and maintaining overall market integrity. They are sufficiently high to allow investment in enough generation so that the expected level of unserved energy does not exceed the reliability standard. The Panel has also considered the case for lowering the market price cap and the cumulative price threshold. We have concluded that the potential benefits from lowering these price caps in terms of possible reduced wholesale prices do not outweigh the long term risks associated with having inadequate investment signals to incentivise demand side capacity or marginal new supply so that the total of generation, demand response and transmission interconnection will meet the reliability standard through the review period.*

The reliability standard (no more than 0.002% expected USE in any region in any financial year) was not reconsidered by the Reliability Panel in its 2018 review. As per the Panel's guidelines for reviewing the reliability standards and settings, the reliability standard remains the same as the one in the previous review, unless the Panel considers that there are material grounds for reassessing it. In making this decision, the Panel considers factors including, but not limited to: any changes made to AEMO's Value of Customer Reliability (VCR) measure; and any material changes in the way consumers use electricity, particularly through the use of new technology, which would suggest that some consumers may place a lower value on a

6 AEMO (2018), *The NEM Reliability Framework, Additional information from AEMO to support its Enhanced RERT rule change proposal*, November 2018, p. 10.

7 Reliability Panel (2016), *Review of reliability standard and settings guidelines, Final Determination*, December 2016, pp. 27, 53.

8 Reliability Panel (2018), *Reliability standard and settings review 2018, Final Report*, April 2018, p. iii.

reliable supply of electricity from the NEM.[9, 10, 11] The connection between the reliability standard and the current estimate of aggregate NEM-wide VCR ($33,460/MWh) is not clear.[12] When the Reliability Panel last reconsidered the reliability standard, it said that "modelling indicated that the current reliability standard of a maximum permissible USE of 0.002 per cent would be economically efficient if VCR was assumed to be $30,000/MWh".[13] The reliability outcome is efficient if the MPC is equal to the VCR of the customers who would be interrupted, since consumers are then indifferent between the costs of an interruption and the costs of paying for additional generation at the MPC (ie, since additional reliability would cost more than customers would be willing to pay for it, providing them with additional reliability would make consumers worse off).[14]

In addition to the market mechanism outlined above (which includes in-market reserves), there is also a "strategic reserve" mechanism (the RERT) that provides out-of-market reserves. Through the Projected Assessment of System Adequacy (PASA), AEMO prepares a forecast of expected unserved energy for the upcoming two years, and continuously updates these forecasts. If the AEMO forecast shows that the amount of unserved energy is expected to be greater than 0.002%, such that the reliability standard is expected to be breached, AEMO will procure reserve capacity. The reserve capacity, which can include demand response, is out of

---

9 AEMC (2016), *Review of Reliability Standard and Settings Guidelines, Final Guidelines*, December 2016, p. 5.

10 The reliability standard has been 0.002% expected USE since the start of the NEM in 1998 (Reliability Panel (2014), *Reliability Standard and Reliability Settings Review 2014, Final Report*, July 2014, p. 24). However, in the past the standard has been defined as a long-term average, such that exceeding 0.002% in any one year was not have been considered a breach of the standard (the outages in Victoria and South Australia in 2008/9 resulted in USE of 0.004% in Victoria and 0.003% in South Australia, but the reliability standard was judged not to have been breached (see NEMMCO (2009), *Power System Incident Report – Actual Lack of Reserve (LOR3) in Victoria and South Australia Regions on 29-30 January 2009*, May 2009, p. 4; Reliability Panel (2009), *Annual Market Performance Review 2008-09, Draft Report*, November 2009, p. 3). The standard is now clearly defined as 0.002% expected USE per region per year.

11 Note that in other jurisdictions VCR is often called the Value of Lost Load (VOLL). These two labels refer to the same underlying concept.

12 $33.46/kWh represents the aggregate NEM-wide VCR estimate, across residential, business, and direct connect customers. See AEMO (2014), *Value of Customer Reliability Review, Final Report*, September 2014, p. 2.

13 Reliability Panel (2014), *Reliability standard and settings review 2018, Final Report*, July 2014, p. 25.

14 See also discussion in AEMC (2013), *Advice to SCER on linking the reliability standard and reliability settings with VCR, Final report*, December 2013, section 2.2. See also William Hogan (2005), "On an 'Energy Only' Electricity Market Design for Resource Adequacy", *Working Paper*, pp. 10-11; Steven Stoft (2002), "Power System Economics – Designing Markets for Electricity," *Wiley-IEEE Press*, 1, pp. 149-150; and Samuel Newell et al. (2012), "ERCOT Investment Incentives and Resource Adequacy," Prepared for ERCOT, p. 12.

the market and can be dispatched by AEMO in periods where reserves are low and the market is expected to fail to clear (and where load shedding would otherwise be needed).[15]

"Expected" USE in AEMO's forecast refers to the fact that AEMO's forecasting captures the inherent uncertainty in future outcomes in a probabilistic sense: there are many "likely" outcomes with little or no USE and other less likely outcomes in which the level of USE is higher. The forecast of "expected" USE is the weighted average USE across all the scenarios, where the weights are the probabilities that each outcome will eventuate, according to AEMO's models. The nature of the electricity system is such that the most likely outcomes have USE below the expected level, but there is a "tail" of unlikely outcomes with USE at or above the expected level. Some of the "tail" outcomes have much larger amounts of USE, but are very unlikely and therefore contribute relatively little to the expected level of USE.[16] Outcomes in the tail are referred to as "high-impact, low-probability events" or HILP events. These events could, for example, result from several generators breaking down simultaneously, at the same time as demand is particularly high due to hot weather.[17] These events do not include cascading system-wide outages, since such outages result from network or security faults rather than wholesale-level reliability.

AEMO can use the RERT if its forecast shows that USE is expected to exceed 0.002%. If the forecast amount of USE is less than this quantity, AEMO cannot contract for reserves, and the market mechanism alone operates to deliver wholesale-level reliability.[18] For a projected shortfall in reserves where AEMO has less than seven days of notice (short-notice RERT), AEMO operationalises the reliability standard through the Lack of Reserve (LOR) declarations.[19]

We understand that AEMO has concerns over the RERT. In particular, AEMO is concerned that in circumstances where its modelling shows expected USE of less than 0.002%, there may nonetheless be significant risks to wholesale-level reliability such that it would be beneficial to make use of the RERT mechanism even though the reliability standard is not expected to be breached. AEMO is concerned about HILP events which contribute only a small amount of expected USE, and which therefore do not trigger the availability of the RERT, but which

---

15 AEMO (2018), *The NEM Reliability Framework, Additional information from AEMO to support its Enhanced RERT rule change proposal*, November 2018, pp. 4-5.

16 AEMO (2018), *The NEM Reliability Framework, Additional information from AEMO to support its Enhanced RERT rule change proposal*, November 2018, p. 20.

17 AEMO (2018), *The NEM Reliability Framework, Additional information from AEMO to support its Enhanced RERT rule change proposal*, November 2018, p. 14.

18 The National Electricity Rules currently allows AEMO discretion to determine how the reliability standard is met. This could be interpreted as AEMO having the ability to procure reserves in case of a failure in the future, even if the reliability standard is currently met. See AEMC (2018), *Enhancement to the Reliability and Emergency Reserve Trader, Options Paper*, October 2018, p. 3.

19 Although there is no explicit link between the LOR framework and the reliability standard, AEMO assumes that if a LOR is identified, there is a risk that the 0.002% expected USE standard is also exceeded. See AEMC (2018), *Enhancement to the Reliability and Emergency Reserve Trader, Options Paper*, October 2018, pp.10, 25.

would have significant impacts on consumers if they happen. As a result AEMO has proposed a rule change to make the RERT available in circumstances where expected USE is not greater than 0.002%.[20] The AEMC is currently assessing AEMO's proposed rule change.

B. AEMO's rule change proposal

In March 2018, AEMO submitted two rule change requests to the AEMC related to the RERT. One of the proposed changes related to the reinstatement of Long Notice (10+ weeks) RERT provisions prior to the summer of 2018/2019. The other one (which we are focused on in this report) relates to a more comprehensive suite of reforms to the RERT framework, "taking into account a broader risk assessment framework when procuring the RERT and the standardization of RERT products".[21] Specifically, AEMO recommended that the RERT procurement should be delinked from the reliability standard, and the rules for procuring RERT should take into account:

- The nature of tail risk,[22] via the use of a range of supplementary metrics in addition to expected USE
- The risk appetite of consumers for different levels of load shedding, and
- The cost structure and optimal mix of resources to prevent or mitigate load shedding

In its rule change proposal and subsequent information provided in support of the rule change proposal, AEMO outlines three high-level concerns with the current NEM reliability framework which are the foundation for AEMO's proposed reforms.[23]

AEMO's first concern is that the current reliability framework does not adequately describe the shape and severity of tail risk. AEMO believes that tail risks are increasing due to multiple factors, including (1) rising temperatures due to climate change, (2) increased intermittent wind and solar generation on the NEM system, and (3) a tightening of the supply-demand balance, following significant retirements of thermal generation. AEMO notes that these factors may increase the amount of USE in a non-linear fashion, and may also increase the uncertainty in USE leading to a wider range of potential USE outcomes. Expected USE is calculated as an average across simulations, weighted by the probability of three different levels of maximum demand occurring.[24] Expected USE will change if the probability of a different level of maximum demand changes. However, the probability that any level of maximum demand will occur is not known with certainty, but has to be estimated. This estimate also has a probability distribution, a measure of how certain we are that our estimates are likely to be

---

[20] AEMO (2018), *Reliability and Reserve Trader Rule change proposals*, March 2018.

[21] AEMC (2018), *Enhancement to the Reliability and Emergency Reserve Trader, Consultation Paper*, June 2018.

[22] ie, the risk of extreme events which have high impacts, but are very unlikely to occur.

[23] AEMO (2018), *The NEM Reliability Framework, Additional information from AEMO to support its Enhanced RERT rule change proposal*, November 2018.

[24] AEMO (2018), *The NEM Reliability Framework, Additional information from AEMO to support its Enhanced RERT rule change proposal*, November 2018, p. 13.

accurate. Expected USE will not change when the certainty over the USE probability estimates changes.

AEMO's second concern is that the current reliability framework assumes that all MWh lost in USE events should be equally weighted and does not account for the potentially higher costs per MWh of longer-duration USE events. AEMO establishes and periodically updates an estimate of VCR based on customer surveys. However, surveys suggest VCR varies by customer segment, by the timing of when USE events occur, and by the duration of USE events. AEMO also questions the accuracy with which survey respondents reflect their true VCR due to the rarity of USE events. AEMO recommends that instead of only using VCR, AEMO should also consider customers' tolerance for load shedding in terms of maximum acceptable duration and scale for a USE event.

AEMO's final concern is that the reliability framework ignores customer risk aversion, which, it says, runs counter to most evidence of human behaviour. By focusing on average USE, the framework assumes customers are risk neutral to outages, regardless of magnitude and duration. However, according to AEMO, economic theory and the prevalence of insurance products demonstrate that customers are willing to pay a premium in order to mitigate the risk of extreme events. AEMO concludes that the RERT should be viewed as a form of insurance to manage tail risks, and that therefore the current arrangements are inadequate since they focus on expected USE.

C. Purpose of this report

The Brattle Group was engaged by the AEMC to provide advice on the economic theories associated with the concepts of loss aversion and risk aversion, and how these theories and concepts may be relevant to the framework for wholesale-level reliability in the National Electricity Market (NEM).

AEMO forecasts the expected quantity of USE for the purpose of determining whether the reliability standard will be met. If the standard will not be met, AEMO is able to procure reserves for the RERT mechanism. In preparing its forecast of USE, AEMO considers a wide range of different events which, if they were to happen, would result in USE. Of these events, the ones which are most likely to occur would result in relatively small amounts of USE because the events would be managed by shedding a small amount of load for a short period of time. However, other events included within AEMO's modelling would result in larger impacts (if those events were to happen), because managing the events would require significant amounts of load to be shed, and/or load shedding for extended periods.

Evaluating the impact of reliability events is non-trivial. Impacts may vary non-linearly with the duration, timing, location and extent of the event. For example: an event that lasts for 5 hours might be more than 10 times as "bad" as an event that lasts 30 minutes; an event that cuts off 100,000 people might be more than 10 times as "bad" as an event that cuts off 10,000 people; an event that occurs during a popular auspicious sporting event may be worse than if the same event occurred under similar conditions the next day. One approach to evaluating non-linearity in the impact of HILP events would be to assess whether VCR, expressed in $/MWh, varies across a range including these more extreme events. This approach has been recommended by the AER to address HILP events in the context of transmission planning, where the AER has advised that proponents of transmission investment value any outages with a VCR that is, "appropriate to the range and duration of customers that the HILP event would affect".[25] The AER further states they are concerned that any approach that uses alternative probability weights for HILP events will lack transparency.[26]

Under the current framework for managing wholesale-level reliability and determining whether the reliability standard will be met, the impacts of events are estimated in terms of USE (the quantity of energy not supplied if those events were to happen). Implicitly, therefore, within this framework is an assumption that the impacts of load shedding are proportional to the amount of USE. The Reliability Panel has explained that this makes sense because, in the event of a lack of wholesale-level reliability, the consequential mismatch between supply and demand would be managed by "rotating" outages. Thus the experience of a customer who is interrupted is not strongly dependent on how many other customers are also interrupted.[27] In the event of load-shedding, load is to be shed equitably between NEM regions, and customers are disconnected following a priority ordering. The load of residential customers tends to be shed first.[28] The Reliability Panel has several times considered standards that could be expressed in other ways, such as the probability of a loss of load event (LOLP),[29] but has chosen to continue using a reliability standard expressed in terms of USE.

In this report we will not examine whether current estimates of VCR and the use of the USE metric are sufficiently capturing the impact of HILP events, but rather on how to weight the impacts of these events if they occur. The current approach estimates expected USE by multiplying the estimated USE for each event, if that event happens, by the probability of it happening, summed over all events. Other approaches may put more weight on HILP events. Theories of consumer behaviour such as loss aversion and risk aversion may motivate such alternative weightings. It is worth noting that consumers only face risks in relation to wholesale-level reliability because there is a (large) difference between the price that retail customers pay for an additional unit of electricity and the value they obtain from having that unit supplied. If the two were equal, then customers would be indifferent between purchasing

---

25 AER, *Application guidelines for the regulatory investment tests, Final decision*, December 2018, p. 30.

26 AER, *Application guidelines for the regulatory investment tests, Final decision*, December 2018, p. 31

27 The Reliability Panel said: "The Panel considered the possibility of introducing a hybrid standard in 1998.[footnote omitted] At the time, the Panel recognised that, in general, energy shortfalls to individual consumers would be managed by rotating the shortfalls. As a result, for all probable incidences of shortfall due to reliability, individual consumers would experience very similar effects regardless of how many others were also affected." (Reliability Panel (2007), *Comprehensive Reliability Review, Final Report*, December 2007, p. 24.)

28 AEMC (2013), *Advice to SCER on linking the reliability standard and reliability settings with VCR, Final report*, December 2013, p. 18.

29 Reliability Panel (2014), *Reliability Standard and Reliability Settings Review 2014, Final Report*, July 2014, Appendix C.

electricity and being interrupted, and no issues of risk or loss aversion would arise.[30] Although the wholesale price will be very high (at the MPC) during a wholesale-level reliability event, retail prices generally do not pass through changes in the wholesale price. Retailers will benefit, during a wholesale-reliability event, since they no longer need to purchase power at the MPC for their customers who are interrupted. If the difference between the avoided costs of purchasing wholesale power at the MPC and the lost retail revenue were passed through to consumers, this would reduce (or eliminate) the wholesale-level reliability risks that consumers face.[31]

This report considers only circumstances in which a lack of wholesale-level reliability could lead to outages, and does not consider other events connected with networks or system security. As discussed above, when the system operates as it is supposed to, an imbalance between supply and demand that cannot be cleared by the price mechanism or the RERT will be resolved by load shedding and rotating outages. Such outages are qualitatively different from the more widespread and extended outages that can be caused by system security problems. Cascading outages and, in the limit, "system black" events have much larger impacts because they affect a wider geographic area and last longer. While our report focuses only on wholesale-level reliability, it is important to note that the conditions in which load shedding for wholesale-level reliability reasons are conditions in which the system is likely to be under stress, and outages for other reasons including security reasons may be more likely under those conditions.[32] Furthermore, system operators usually do not have much experience of operating the system under these conditions.[33] Some aspects of reliability framework design, and/or actions by system operators, can reduce outage risk both in relation to wholesale-level reliability and in relation to system security.

In our report we have performed a literature review of these theories in behavioural economics. We have also looked for practical applications of such theories in policy-making in various sectors. In addition, we have reviewed the frameworks for delivering reliability in several wholesale electricity markets outside Australia. The report is structured as follows. Section II introduces and defines loss aversion and risk aversion theories as they have been used in behavioural economics. In this section, we examine academic studies that present different implications for risk mitigation (ie, explanations and analyses to explain their various implications). Section III reviews empirical studies to assess the impacts of these theories in the insurance sector. Appendices A and B examine examples of risk aversion in climate risk analysis and flood defence, respectively. Specifically, we focus on how risk aversion theories

---

30 This is true for the average customer, for individual customers, their personal VCR may differ from the average.

31 This is in line with the Load Shedding Compensation Mechanism discussed by the AEMC in relation to compensation for load shedding due to reliability reasons. See AEMC (2018), *Wholesale Demand Response Mechanism, Consultation Paper*, November 2018, Appendix D.

32 AEMO had previously proposed an "operational" reliability standard to manage power system reliability during extreme conditions. See AEMO (2018), *AEMO observations: Operational and market challenges to reliability and security in the NEM*, March 2018.

33 We understand that the NEM has had two wholesale-level reliability events leading to load shedding in the last 10 years. See Reliability Panel (2018), *Annual market performance review 2017, Final report*, March 2018, p. 49.

apply to the practical application of cost-benefit tests, estimating high impact, low probability events, and any examples of international jurisdictions that incorporate risk aversion theory into their regulated decision making processes. Section IV looks at the electricity sector specifically, and reviews international jurisdictions to see if they explicitly (or implicitly) account for loss aversion theory and HILP in resource planning and operations. Finally, Section V draws on the lessons learned and concludes with recommendations for the NEM, taking into account its unique circumstances.

II. Consumer risk preferences

A. Introduction to risk aversion

The fields of economics and finance have long recognized that individuals often value uncertain or risky outcomes differently from a pure expected value approach. This was first recognized by Swiss mathematician Daniel Bernoulli in 1738, who argued that two people, facing the same lottery, may value it differently because of a difference in their psychology.[34] This challenged the predominant paradigm at the time (put forward by Fermat and Pascal among others), that the value of a lottery should be equal to its mathematical expectation (the sum of probability weighted outcomes) and hence identical for all people, independent of their risk attitude.[35]

In general, economists have observed that most people would prefer a less risky outcome over a riskier alternative of equivalent expected value.[36&37] For example, if given the choice between $50 and a "fair bet" to flip a coin and win either $100 or nothing, most people would choose the former over the latter in most circumstances, even though both have an *expected* outcome of $50.[38] Because people prefer the certain outcome, in some situations they have to be offered a "premium" to switch to the risky outcome: for example, $110 or nothing, instead of a certain $50. The intuition that people are "risk-averse" explains many real world applications such as the demand for insurance products and the risk-return trade-off observed in the pricing of financial assets.[39] Insurance products such as health insurance, property insurance, unemployment insurance, flood insurance and so forth, help consumers reduce their risks by spreading those risks over a wide pool of consumers. Like insurance products, financial markets

---

34 Louis Eeckhoudt, Christian Gollier and Harris Schlesinger (2005), "Economic and Financial Decisions under Risk," *Princeton University Press*, pp. 3-5.

35 Louis Eeckhoudt, Christian Gollier and Harris Schlesinger (2005), "Economic and Financial Decisions under Risk," *Princeton University Press*, pp. 3-5.

36 Ted O'Donoghue and Jason Somerville (2018), "Modeling Risk Aversion in Economics," *Journal of Economic Perspectives,* 32(2), pp. 91-114.

37 Charles Holt and Susan Laury (2002), "Risk Aversion and Incentive Effects," *The American Economic Review*, 92(5), pp. 1644-1655.

38 Matthew Rabin and Richard Thaler (2001), "Anomalies: Risk Aversion," *Journal of Economic Perspectives,* 15(1), p. 226.

39 Ted O'Donoghue and Jason Somerville (2018), "Modeling Risk Aversion in Economics," *Journal of Economic Perspectives,* 32(2), pp. 91-114.

help investors to reduce their overall financial risks by providing them with an opportunity to hold a diverse portfolio of securities.[40]

The "standard" approach in economics to explaining why people are risk averse has been the theory of expected utility. For many economists expected utility is synonymous with risk aversion. [41] However, empirical findings in experimental economics and psychology have shown that expected utility theory is not sufficient to explain many of the actual decisions made by real people over risky alternatives. Consequently, several new theories explaining risk aversion have arisen, most prominently loss aversion. We discuss expected utility theory, some of its shortcomings, loss aversion and the broader theory surrounding it in more detail below. Each of these theories has the same broad prediction that consumers dislike risk, but will differ in some of their more nuanced implications.

Takeaways for the NEM

- Economists have observed that in many settings individuals do not behave as though they are risk neutral: rather, most people seem to prefer certain to uncertain outcomes—ie, they are risk averse.
- In many circumstances, people do not behave as though they think about uncertain outcomes using pure probability weights.
- It therefore makes sense to consider whether risk aversion has implications for policy in relation to risk and uncertainty in the NEM's reliability framework.

B. Expected utility theory

Expected utility theory builds on one of the foundational building blocks of economic theory - diminishing marginal utility. Utility is an abstract term used by economists to measure the level of satisfaction derived from an activity. Diminishing marginal utility means that each successive unit of an item gives the consumer less satisfaction than the previous unit. For example, a consumer will get more happiness on a hot day from their first ice-cream than their

---

40 Financial asset pricing models such as the well-known Capital Asset Pricing Model (CAPM) assume that investors can eliminate all idiosyncratic risk by holding a diverse portfolio of assets. Consequently, any risk premium that they receive in return for holding risky assets only applies to the residual "systematic" risk, which cannot be diversified away. For the rest of this paper, we do not distinguish between systematic and idiosyncratic risk.

41 This has created some confusion in the economic literature, with risk aversion and expected utility theory being used interchangeably. Consequently, alternative theories describing preferences to reduce risk have been described as alternatives to the theory of risk aversion, whereas they would be better described as alternatives to expected utility theory. This has led to a divergence between the common meaning of risk aversion – ie, demonstrating a preference to reduce risk - and the economic definition, which has been limited to expected utility theory. Following O'Donoghue and Somerville (2018) we simplify this discussion by labelling risk aversion as a preference to reduce risk and separating it from expected utility, which is one of several theoretical models that economist use to explain risk aversion.

second, and so on. This also applies to wealth, where successive units of wealth (or income) are valued less than the previous unit (the underlying rationale is the same since wealth can be exchanged for ice-creams or any other good).[42] The theory of expected utility would explain the individual's distaste for the fair bet described above as being driven by the diminishing marginal utility of wealth. In this interpretation, the individual has a certain gain of $50 which they can either keep, or risk in a coin flip to win a second $50 or lose the first $50. According to the theory of diminishing marginal utility, the second $50 is worth less than the first $50, so the expected *utility* of the fair bet is negative, even though the expected change in wealth is zero. Thus the risk averse individual keeps the first $50 and does not risk it in a fair bet.

The degree of risk aversion can be measured by the size of the discount that an individual is willing to take to avoid the risky choice. In modern portfolio theory, this is defined as the additional expected reward that an investor requires to take on additional risk.[43] Other definitions of risk aversion are specific to expected utility theory and are used to create mathematically tractable utility functions that economists can use in their modelling. For example constant *absolute* risk aversion says that when an individual receives an increase in wealth, they will not change the *magnitude* of their investments in risky assets, whereas constant *relative* risk aversion says that they will not change the *share* of wealth invested in risky assets. Both measures can be increasing or decreasing rather than constant. For example decreasing relative aversion says that as wealth increases, individuals will hold a larger share of investments in risky assets. Generally, economists prefer relative rather than absolute measures since it makes it easier to compare sensitivities. There is no particular reason to model expected utility with increasing or decreasing risk aversion, since there are no definitive empirical results, or universally appealing theoretical properties supporting either choice.[44]

As a model of consumer behaviour, expected utility theory has a major flaw – it cannot plausibly account for consumer preferences over small to moderate risks. For most adults in industrialized countries, an additional $100 will have practically no impact on lifetime wealth. Thus, the rate at which marginal utility of wealth diminishes as wealth increases should be unnoticeable for changes in wealth like $100. Thus, individuals should be indifferent between taking the coin flip for either $100 or $0 (mentioned earlier) and the choice of getting $50 with certainty. Yet this type of fair bet over small stakes is routinely rejected in laboratory and real world experiments. Mechanically, the fact that consumers reject small stakes gambles implies that their marginal utility of wealth is decreasing rapidly. Thus, if the theory of diminishing marginal utility explains risk aversion for small risks, the theory implies that individuals will completely reject larger risks, irrespective of what compensation is offered for bearing the

---

[42] Samuel Bowles (2006), "Microeconomics, Behavior, Institutions, and Evolution," *Princeton University Press*, pp. 102-103.

[43] This approach is actually just an approximation of the risk premium and will hold true for small changes in risk. For larger changes in risk, the risk premium will also depend upon the other moments of the distribution of the risk. See Louis Eeckhoudt, Christian Gollier, & Harris Schlesinger (2005), "Economic and Financial Decisions under Risk", *Princeton University Press*, pp. 11-12.

[44] Louis Eeckhoudt, Christian Gollier, & Harris Schlesinger (2005), "Economic and Financial Decisions under Risk", *Princeton University Press*, pp. 17-18.

larger risk.[45] Rabin (2000) shows that this holds true for any functional form of the utility function and context (wealth level and assets under risk).[46] For example, Rabin (2000) shows that an expected utility maximizer who rejects a 50-50 bet to lose $100 or gain $200 will turn down a 50-50 bet to lose $200 or gain $20,000. Intuitively, this degree of risk aversion seems implausible, since the larger bet has an expected payoff of $9,900 and a maximum loss of $200.[47]

Over larger stakes, expected utility explains risk averse preferences and has many applications in economics and finance since it allows for the easy quantification and comparison of risk aversion between individuals.[48] Although alternative specifications for expected utility have arisen to deal with the "calibration" issue between small and large risk preferences, [49] there are still other aspects of consumer behaviour that are better explained by alternative theories of risk aversion.[50]

Takeaways for the NEM

- The theory of expected utility is widely used in economics and finance and across multiple sectors.
- Expected utility theory assumes that risk aversion arises because the incremental value of additional wealth decreases as an individual gets richer. This means that money potentially lost is valued higher than money potentially gained. Expected utility theory explains why individuals are risk averse with respect to significant risks (large impacts and non-negligible probabilities), but it implies that individuals should not be risk averse with respect to risks that are small relative to total wealth.

---

45 Matthew Rabin and Richard Thaler (2001), "Anomalies: Risk Aversion", *Journal of Economic Perspectives,* 15(1), pp. 221-222.

46 Matthew Rabin and Richard Thaler (2001), "Anomalies: Risk Aversion", *Journal of Economic Perspectives,* 15(1), pp. 221-222.

47 Matthew Rabin and Richard Thaler (2001): "Anomalies: Risk Aversion", *Journal of Economic Perspectives,* 15(1), p. 224.

48 Ted O'Donoghue and Jason Somerville (2018), "Modeling Risk Aversion in Economics", *Journal of Economic Perspectives,* 32(2), p. 92.

49 For example, Schechter (2007), finds that modelling expected utility on diminishing marginal utility of income rather than wealth, provides reasonable expected utility parameter estimates for small bets. See Laura Schechter (2007), "Risk aversion and expected-utility theory: A calibration exercise", *Journal of Risk and Uncertainty,* 35(1), pp. 67-76.

50 Ted O'Donoghue and Jason Somerville (2018), "Modeling Risk Aversion in Economics", *Journal of Economic Perspectives,* 32(2), p. 98.

C. Loss aversion & prospect theory

Prospect theory was first introduced in 1979 by two psychologists, Daniel Kahneman and Amos Tversky, to capture the key features of a number of experiments where decisions made under risk violated the predictions of expected utility theory.[51] Daniel Kahneman was awarded the 2002 Nobel Prize in economics "for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty".[52]

Within this broader model of consumer decision making under risk is the concept of loss aversion, which states that individuals are more sensitive to losses than equivalent gains, irrespective of their wealth or the magnitude of the losses.[53] Although loss aversion is just one of several components that constitutes prospect theory, the two concepts are often used interchangeably in practice. However, we will discuss them separately, since one of the other components of prospect theory, probability weighting, can act as alternative explanations for risk aversion independent of loss aversion.

Loss aversion suggests that decisions are made by evaluating gains and losses relative to a reference point, and that gains are valued less than equivalent losses. Reference points can be decision- and circumstance-specific. Oftentimes the status quo is assumed to be the reference point.[54] In a consumer survey to determine the value of lost load for customers in California, Hartman, Doane and Woo (1991) found that preferences for reliable electric supply were influenced by the current ("status quo") level of reliability the customer faced.[55] In particular, the authors split their survey respondents into two groups – a high reliability group, who had experienced approximately 3 outages of two-hours each in the past year, and a low reliability group, who had experienced 15 outages of four-hours each in the past year. Electricity prices for the high reliability group were about 30 percent higher than the low reliability group.[56]

---

51 Although first introduced in 1979, prospect theory was only generalized by Kahneman and Tversky in 1992 under the label "cumulative prospect theory". For the most part economists are referring to the latter theory when they refer to prospect theory. See Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), pp. 173-175.

52 The prize was shared with experimental economics pioneer Vernon Smith. Amos Tverskey had already passed away when the prize was awarded. See The Nobel Prize (2002), *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2002*, October 2002.

53 Daniel Kahneman and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), pp. 263 – 292. See also Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), p. 175.

54 Samuel Bowles (2006), "Microeconomics, Behavior, Institutions, and Evolution," *Princeton University Press*, pp. 107-108.

55 Note that reliability of supply in this context means any outage, not just generation adequacy.

56 Raymond Hartman, Michael Doane, and Chi-Keung Woo (1991), "Consumer Rationality and the Status Quo," *Quarterly Journal of Economics,* 106(1), p. 149.

Study participants were asked to rank their preferences over six different combinations of reliability and electricity prices, including the status quo. Despite significant differences in outage rates and electric prices between the two groups, participants in both groups preferred their *own* status quo, with about 60% of customers in each group selecting it as their most preferred option.[57] Kőszegi and Rabin (2006, 2007, 2009) hypothesize that consumers form expectations or beliefs about the near future, and evaluate prospective gains and losses relative to that reference point.[58] This is consistent with the idea of using the status quo as the reference point, if the near future is expected to remain unchanged. In the coin flip example previously discussed, if an individual were to take the offer of $50 with certainty as their expected outcome, they would see the coin flip for $100 or $0, as a gain of $50 or a loss of $50 relative to their reference point. Under loss aversion, they would "feel" the loss of $50 more than the gain of $50, and the expected value would be negative, so they would reject the bet. This is similar to expected utility, except that under expected utility the result was driven by the diminishing marginal utility of wealth, whereas under loss aversion, there is direct dissatisfaction from the loss itself.

Taken alone, the concept of loss aversion simply states that losses are valued more than gains. In practice, loss aversion is often modelled so that there is diminishing sensitivity in the value function, such that the relative difference in value between a small loss and an equally small gain is larger than the relative difference between a large loss and an equally large gain.[59] Increasing sensitivity means that additional gains are positively valued, but at a decreasing rate, while additional losses are negatively valued, but also at a decreasing rate. This implies that when making decisions that are over pure gains, relative to the reference point, people are risk averse, but when decisions are over pure losses, they are risk seeking.[60] For example if the reference point is zero, an individual would prefer gaining $50 with certainty to a coin flip for $100 or zero. However, if they were faced with the gamble framed as a loss, they would rather choose the coin flip than pay the $50 with certainty. This phenomenon is called the reflection effect and based on empirical observations by Kahneman and Tversky.[61] An experiment conducted with Brisbane homeowners, after the 2011 Brisbane floods, found evidence to support the reflection effect, with individuals who had suffered losses from the flood, being 50% more likely to opt for a risky gamble (choosing a low probability, high stakes scratch card over a certain $10) than their unaffected neighbours.[62]

---

57 Daniel Kahneman, Jack Knetsch, and Richard Thaler (1991), "The Endowment Effect, Loss Aversion, and Status Quo Bias", *Journal of Economic Perspectives*, 5(1), p. 198.

58 Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), p. 179.

59 Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), p. 175.

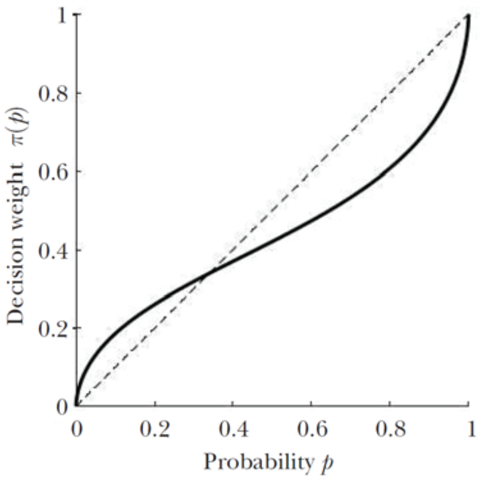
60 Daniel Kahneman and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), pp. 263-292.

61 Daniel Kahneman and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), pp. 268-269.

62 Lionel Page, David Savage, and Benno Torgler (2014), "Variation in risk seeking behaviour following large losses: A natural experiment", *European Economic Review*, 71, pp. 121-131.

The probability weighting function is another component of prospect theory and is often coupled with or thought to be part of loss aversion. Put simply, this concept states that individuals may make decisions based on subjective decision weights that differ from the underlying objective probability in systematic ways. Based on experimental results, Tversky and Kahneman (1992) proposed a probability weighting function that overweighs low probabilities and underweights high probabilities.[63] These transformed probabilities do not represent erroneous beliefs, and are based on objective probability.[64] For example, O'Donoghue and Somerville (2018) provide the following scenario: consider a binary gamble with an objective probability p of receiving $10, and an objective probability 1-p of receiving $100. For small values of p (such as 0.2), the $100 outcome receives a decision weight smaller than its probability while the $10 outcome receives a decision weight that is larger. When comparing to an expected utility maximiser, this gamble will be perceived as less attractive and therefore the probability weighting generates a source of risk aversion. Conversely, in a scenario with large values of p (such as 0.8), the opposite holds, where the $100 outcome receives a decision weight larger than its probability and the $10 outcome receives a decision weight smaller, thus generating a source of risk seeking.[65] Figure 1 shows an example of a probability weighting function. Actual probability is shown by the 45 degree line. Relative to this, decision weights overweight low probability events and underweight high probability events.

Figure 1: Example of a probability weighting function



Sources: Ted O'Donoghue and Jason Somerville (2018), "Modelling Risk Aversion in Economics", *Journal of Economic Perspectives*, 32(2), p. 101.

Loss aversion explains many observed behaviours that were inexplicable using expected utility theory, ranging from the existence of insurance for small expected risks (such as failure of white goods) to the equity premium puzzle. Extended warranties for electronics or appliances (termed white goods) act as an insurance policy to the consumers – if the purchased device or equipment

---

63 Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), p. 176.

64 Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), pp. 177-78.

65 Ted O'Donoghue and Jason Somerville (2018), "Modelling Risk Aversion in Economics", *Journal of Economic Perspectives*, 32(2), pp. 100-102.

were to fail, the customer would get some form of guaranteed payment or product replacement. Given that appliance failure would be inconsequential to lifetime wealth, expected utility theory was unable to explain the popularity of these insurance products. Moreover, the high markups that were common on appliance insurance implied unrealistically high levels of risk aversion using expected utility theory.[66] For example, Jindal (2014) found that while the profit margin on electronics or appliances typically range from 15% to 20%, the profit was more than 200% on extended warranties (this suggests the risk premium on such warranties is much higher than the actuarially fair value).[67] Through surveying 550 consumers who had purchased washing machines, the author further identified that consumers were viewing repairs as losses, and that loss aversion and peace of mind were the key reasons they bought extended warranties.[68,69]

The equity premium puzzle refers to the fact that the average return on equities has historically exceeded the return on risk-free government bonds by a much higher margin than any asset pricing models predicted. Returns in asset pricing models are predicated on rewarding the additional risk that equity holders face, compared to bond holders, based on expected utility theory. Expected utility theory predicts that the equity premium should be much smaller than what is observed. Benartzi and Thaler (1996) propose an alternative explanation based on loss aversion that better explains the magnitude of the risk premium for holding equities. They hypothesize that investors think about short-term volatility, with frequent losses and gains, and that this distracts them from the prospect of long-run returns, so that they tend to require a larger equity premium as compensation for the short-term volatility.[70] Benartzi and Thaler term this short-term assessment of losses "myopic loss aversion".

Empirically, loss aversion can be difficult to demonstrate outside of laboratory settings since reference points are unobserved. However, a number of empirical findings have emerged that support the loss aversion hypothesis. For example Odean (1998) found that investors were more likely to sell stocks that were experiencing gains than those experiencing losses,[71] while Frazzini (2006) found that investors were more likely to sell stocks after receiving positive news

---

66 Matthew Rabin and Richard Thaler (2001), "Anomalies: Risk Aversion," *Journal of Economic Perspectives,* 15(1), pp. 226-227.

67 Pranav Jindal (2014), Risk Preferences and Demand Drivers of Extended Warranties, *Marketing Science*, 34(1). See also EurekAlert (2015), *Selling extended warranties via independent companies lowers price but hurts consumers: INFORMS*, January 2015.

68 While for the most part none of the studies referenced are Australia specific, we think that these observations would hold in Australia.

69 Jindal (2014) estimated that consumers felt the pain of repairs 3-4 times more than paying for the price of the product itself. See Pranav Jindal (2014), Risk Preferences and Demand Drivers of Extended Warranties, *Marketing Science*, 34(1).

70 Matthew Rabin and Richard Thaler (2001), "Anomalies: Risk Aversion," *Journal of Economic Perspectives,* 15(1), pp. 226-227.

71 Terrance Odean (1998), "Are Investors Reluctant to Realize Their Losses," *The Journal of Finance*, 53(5), pp. 1775-1798.

(pushing the price up) than negative news (pushing the price down).[72] Genesove and Mayer (2001) found similar results in the housing market, where condominium sellers in downtown Boston would set higher prices when facing a loss on their original purchase price, than those sellers facing a potential gain.[73]

Understanding risk aversion through the lens of loss aversion rather than expected utility raises the question of whether loss aversion is a consumer taste, or a behavioural bias in need of correction. If the latter were the case, one could argue that loss aversion would lead to outcomes that individuals might later regret, and that if they were better informed and were made aware of this bias, they would take measures to correct it. Behavioural neuro-economist Colin Camerer (2005) suggests that "disliking potential loss is a legitimate preference if, when loss is felt, there is a substantial hedonic sensation that is either brief but painful or long lasting, but disliking a possible loss is a judgment error if losses are transitory."[74] Camerer defines transitory losses as feelings or perceptions that may be overstated if the decision maker is in a heightened emotional state, which does not reflect their general disposition.

Feelings of regret and a desire for behavioral correction are certainly the case for other behavioral biases, such as hyperbolic discounting. Hyperbolic discounting is a description of how people may discount future periods too strongly, resulting in ongoing procrastination.[75] For example, an individual may wish to go to the gym every other day to get in shape. However, on any given day, they do not feel like going and decide to go tomorrow instead, but when tomorrow becomes today, they make the same choice and never end up going. When informed of their behavioral bias, these individuals may take actions to correct for it. For example, they might create a commitment device, like meeting a friend at the gym, to change their short-term payoff function. There is a large literature on correcting hyperbolic discounting that discusses naïve versus sophisticated (informed) hyperbolic discounters and commitment devices.[76] Dean Karlan, a professor of economics at Yale, even started a service where individuals could commit to a goal, such as going to gym every day, where failure to achieve this goal would result in an automatic donation by the transgressor to charity.[77] This changes the payoff for going to gym on any given day (and uses the concept of loss aversion).

---

72 Andrea Frazzini (2006), "The Disposition Effect and Underreaction to News," *The Journal of Finance,* 61(4), pp. 2017-46.

73 David Genesove and Christopher Mayer (2001), "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *The Quarterly Journal of Economics*, 116(4), pp. 1233-1260.

74 Colin Camerer (2005), "Three Cheers—Psychological, Theoretical, Empirical—for Loss Aversion," *Journal of Marketing Research,* XLII, p. 131.

75 See for example, David Laibson (1997), "Hyperbolic Discounting and Golden Eggs," *Quarterly Journal of Economics*, 112, pp. 443-77.

76 See for example Richard Thaler and Hersh Shefrin (1981), "An Economic Theory of Self-Control," *Journal of Political Economy*, 39, pp. 392-406. See also, Dan Ariely and Klaus Wertenbroch (2002), "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science*, 13, pp. 219-224 and Gharad Bryan, Dean Karlan and Scott Nelson (2010), "Commitment Devices", *Annual Review of Economics*, 2, pp. 671-698.

77 Or if you wanted to be highly motivated, failure could result in money being donated to an anti-charity – an organisation antithetical to your personal belief system.

In contrast to hyperbolic discounting, there does not seem to be a literature seeking to change loss averse behaviour, leading us to believe that loss aversion is mostly regarded as a feature of consumer preferences, rather than a bias in need of correction. The only exception that we have identified is in financial planning. Loss aversion is not wealth maximizing behaviour. In activities where wealth maximization is the explicit goal, such as in relation to investing, loss averse behaviour may be better thought of as a behavioural bias rather than a preference.[78] In the same way that a financial advisor may suggest that a client consider cutting back on vacations to increase savings, they may also try to reframe investment decisions to avoid loss averse behaviours such as holding on to losing stocks.[79] Framing involves changing the decision maker's reference point. How problems are framed matters for the outcome. For example, monetary incentives framed as losses are more effective in motivating workers when compared to framing them as gains.[80]

Takeaways for the NEM

- Loss aversion and prospect theory are an established part of the behavioural economics framework, and explain why individuals display risk aversion in respect of relatively small risks (small impacts and non-negligible probabilities).
- The behavioural economics literature often recommends steps that should be taken to help individuals "correct" some behavioural biases in order to avoid later regret.
- This is mostly not the case for loss aversion, which may be better thought of as a component of consumer preferences, rather than a behavioural bias.

78 There is mixed evidence on whether experienced traders are less loss averse. List (2003) and List (2004) found that experienced traders of sports memorabilia were less likely to have an endowment bias (loss aversion over possessions), than naïve/inexperienced traders, even with goods outside of their area of expertise. However, in a separate study, Haigh and List (2006) found that professional futures and options traders recruited from the Chicago Board of Trade exhibited behaviors that were consistent with myopic loss aversion to a greater extent than with students. See John List (2004), "Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace," *Econometrica*, 72(2), pp. 615-625 and Michael Haigh and John List (2006), "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis," *The Journal of Finance*, LX(1), pp. 523-34

79 Vicki Bogan (n.d.), *Risk Aversion vs. Loss Aversion: What is the Big Difference?*, n.d.

80 Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), p. 190.

III. Examples of risk aversion and managing HILP risks in insurance

Risk aversion is the primary paradigm that underlies the demand for insurance, particularly for individual consumers. However, it is not the only possible motive for seeking to acquire insurance products. Large, widely held corporations (and governments) are often assumed to be risk neutral,[81] since they are able to pool risks across projects and spread them across many shareholders.[82,83] We would expect that these risk-neutral entities would find it more cost-effective to self-fund insurance for themselves, rather than paying a premium to another party for the service. Yet such entities often do purchase regular insurance products from third parties. Various reasons other than risk aversion have been postulated to explain the existence of insurance markets for risk-neutral entities. For example, transaction costs arise in trading and contracting activities where neither agent fully trusts the other. By reducing the possibility of non-performance (for example not having the liquidity to pay for damages to goods), insurance policies reduce the transaction costs of trade. Standard contractual clauses frequently require guarantees and/or insurance.[84] Thus the existence of insurance markets does necessarily imply the existence of risk aversion. Nonetheless, the remainder of our discussion below will focus on the demand for insurance by individual consumers, which is widely thought to be driven by risk averse preferences.

The market for insurance for individual consumers is a major economic activity. In 2016 in Australia, individual consumers spent AUD$100 billion on insurance premiums, constituting around 5.9% of Australia's GDP.[85] Not all of these insurance purchases were voluntary decisions - some types of insurance are mandatory, others are provided by employers, and some are subsidized by the government. However, the market for voluntary, unsubsidized insurance

---

81 Erwann Michel-Kerjan, Paul Raschky, and Howard Kunreuther (2011), "Corporate Demand for Insurance: An Empirical Analysis of the U.S. Market for Catastrophe and Non-Catastrophe Risks," *NBER Working Paper Series*, 17403.

82 Bruno Merz, Florian Elmer, and Annegret Thieken (2009), "Significant of 'high probability/low damage' versus 'low probability/high damage' flood events," *Natural Hazard and Earth Systems Sciences*, 9, pp. 1033-1046.

83 Kenneth Arrow and Robert Lind (1970), "Uncertainty and the Evaluation of Public Investment Decisions," *American Economic Review*, 60, pp. 364-78.

84 Göran Skogh (1998), "Mandatory Insurance", *Encyclopedia of Law and Economics*, Bouckaert, B./de Geest, G. (eds.), Entry No. 2400, pp. 526-530.

85 OECD (2018), *OECD Insurance Statistics 2017*, July 2018, p. 42. See also APRA (2018), *Quarterly life insurance performance statistics*, November 2018. Australia's GDP in 2016 was AUD$1.7 trillion, see Australian Bureau of Statistics (2017), *Australian National Accounts: National Income, Expenditure and Product, Dec 2016*, March 2017.

products is still significant. Life insurance,[86] which offers a payment to the policyholder's family in the case of their death, alone accounted for 55% of Australian's insurance premium expenditure.[87] Although the insurance market is large, demand is not uniform across insurance products and some puzzling patterns of consumption emerge.[88] In particular:

- Insurance for high probability low impact events is relatively popular. Studies have shown the consumers pay large premiums to cover small losses in products such as warranties for white goods, with observed premiums as great as ten times that of the fair actuarial value.[89]
- Insurance for low probability high impact events is relatively unpopular,[90] despite that fact that it is often subsidized for consumers in disaster prone areas. For example, private coverage for the 1971 San Fernando earthquake covered only 6% of the insurable damage.[91] Similarly, only 20% of New York City homeowners had flood insurance coverage when Hurricane Sandy caused large scale flooding in 2012. This was despite prior flood damage from Hurricane Irene in 2011.[92]

The first puzzle arises because the theory of expected utility is unable to explain why anyone would purchase insurance for small items, since the (certain) premium is much higher than the expected loss. However, this behaviour can be explained by the theory of loss aversion, described above. Under loss aversion, small losses are felt sharpest.[93] Furthermore, the relatively high probability of these loss events reinforces their value.

The second puzzle is not explained by loss aversion. Rather, it becomes more inexplicable if one assumes that consumers weight low probability events more than their objective

---

86 Note that life insurance policies in Australia are often bundled as a default choice in mandatory superannuation policies. Consumers may not be aware that the life insurance policy is voluntary and that they can opt out. See for example Alice Uribe (2018), "Young super members unlikely to opt-in to life insurance, premiums set to rise", *Australian Financial Review*, May 18.

87 OECD (2018), *OECD Insurance Statistics 2017*, July 2018, Table 8. See also APRA (2018), *Quarterly life insurance performance statistics*, November 2018.

88 While the studies quoted are international, there is no indication that these puzzles are less likely to hold in Australia.

89 David Cutler and Richard Zeckhauser (2004), "Extending the Theory to Meet the Practice of Insurance," *Brookings-Wharton Papers on Financial Services,* pp. 1-53.

90 Ulrich Schmidt (2016), "Insurance Demand under Prospect Theory: A Graphical Analysis," *The Journal of Risk and Insurance*, 83(1), pp. 77-89.

91 Dan Anderson and Maurice Weinrobe (1986), "Mortgage Default Risks and the 1971 San Fernando Earthquake," *Journal of the American Real Estate & Urban Economics Association*, 14(1), pp. 110-135.

92 Wouter Botzen, Howard Kunreuther, and Erwann Michel-Kerjan (2015), "Divergence between individual perceptions and objective indicators of tail risks: Evidence from floodplain residents in New York City," *Judgement and Decision Making*, 10(4), pp. 365-385

93 Matthew Rabin and Richard Thaler (2001), "Anomalies, Risk Aversion," *Journal of Economic Perspectives*, 15(1), pp. 219-232.

probability (this is one assumption of prospect theory). Below we discuss some of the empirical and theoretical findings concerning insurance for low probability high impact events.

Kunreuther and Pauly (2004) propose a model where transaction costs prove prohibitively high for purchasing insurance for high impact low probability events, such as catastrophes. They postulate that consumers know that the risk of a catastrophe occurring is low, but are uncertain of what the actual level of risk is. The author's model assumes that consumers need to know the probability of a catastrophic event occurring in order to assess the value of the insurance policy relative to its cost. Given that the risk of a catastrophe is low to begin with, consumers may find that the costs of learning about the actual risk profile is prohibitively high compared to their ex-ante assessment of the possible benefit.[94]

An alternative hypothesis proposed by Rashky and Weck-Hanneman is that government assistance may crowd out private insurance for catastrophic events. The authors contend that by the very nature of a catastrophic event, consumers can be relatively confident that the government will step in with assistance. This in turn diminishes the value of holding private catastrophe insurance.[95]

A third hypothesis states that sometimes people conflate low probabilities with zero probability. Urbany, Schmit, and Butler (1989) conducted an experiment where respondents were asked to classify the probabilities of picking certain jelly beans from different buckets of jelly bean mixes. Respondents grouped probabilities that were less than 0.1 under the categories "no chance" and/or "no worry". [96] Similar behaviour was also observed and discussed by McClelland, Schulze, and Coursey (1993), where their laboratory results were found to be similar to field evidence that, for low probability events, people either dismiss the risks or worry too much about them.[97]

All three of these hypotheses could explain why consumers do not generally purchase insurance for low impact high probability events. None of them preclude risk aversion as a potential driver of demand for high impact low probability events. Rather they would say that if risk aversion is relevant, there are transaction costs, alternative incentives, or calculation errors which overwhelm the risk aversion. The last hypothesis that people have trouble differentiating low probabilities from zero does, however, directly contradict prospect theory's assumption that low probability events are weighed heavily more heavily than their objective probability.

---

94 Howard Kunreuther and Mark Pauly (2004), "Neglecting Disaster: Why Don't People Insure Against Large Losses?," *The Journal of Risk and Uncertainty*, 28(1), pp. 5-21.

95 Paul Raschky and Hannelore Weck-Hannemann (2007), "Charity hazard—A real hazard to natural disaster insurance?," *Environmental Hazards,* 7(4), pp. 321-329.

96 Joel Urbany, Joan Schmit, and Daniel Butler (1989), "Insurance Decisions (Or the Lack Thereof) for Low Probability Events," *Advances in Consumer Research*, 16, pp. 535-541.

97 Gary McClelland, William Schulze and Don Coursey (1993), "Insurance for Low-Probability Hazards: A Bimodal Response to Unlikely Events," *Journal of Risk and Uncertainty*, 7(1), pp. 95-116.

We provide two further examples of risk aversion and managing HILP risks in the Appendix: climate risk analysis and flood defence.

Takeaways for the NEM

- Risk aversion (due to utility maximization or loss aversion) are likely to be significant drivers for consumers to purchase insurance. Voluntary insurance purchases constitute a significant share of Australian economic activity.
- Insurance for some high probability, low impact events is popular, despite having large premiums relative to the value of the insured item. This is consistent with the theory of loss aversion.
- Insurance for low probability, high impact events is unpopular, even when premiums are subsidized. This is inconsistent with the theory of risk aversion (and expected utility theory), but may be the result of confounding factors such as high transaction costs, or the crowding out of insurance by government-provided catastrophe relief.

IV. Reliability frameworks in other jurisdictions

In this section we describe how decisions about wholesale-level reliability have been taken in PJM, ISO-NE, ERCOT and Great Britain. In seeking to illustrate whether decisions seem to go beyond a standard cost-benefit assessment to adopt a more risk-averse approach to maintaining reliability, or otherwise assign greater weight to HILP events, we look at both how reliability standards are set and how the mechanisms operate which are put in place to achieve them. In the descriptions below, we distinguish between a) assigning a greater weight to HILP events than would be justified by an assessment of the probability and b) assigning a larger impact to HILP events than implied by a standard value of lost load (VOLL) or VCR figure. VOLL, when expressed in $/MWh terms, may be a function of the severity of an outage, so effectively some policy makers may assume a larger VOLL for severe outages than for smaller ones. The latter approach has been recommended by the AER to address HILP events in the context of transmission planning (see discussion in Section I.C).

Under supply and demand imbalance situations, market prices are technically undefined and must be set administratively, such as setting them equal to a defined price cap. The most efficient approach to setting prices during such load shed events is to set the prices equal to the value of service of the average customer that is curtailed involuntarily.[98] This estimated Value

[98] See William Hogan (2005), "On an 'Energy Only' Electricity Market Design for Resource Adequacy," *Working Paper*, pp. 10-11; Steven Stoft (2002), "Power System Economics – Designing

of Lost Load (VOLL) reflects the price that the curtailed customers would have been willing to pay on average to avoid being interrupted. It thus represents a proxy "demand curve" at which the average curtailed customer will be indifferent between being interrupted and being charged a high price for consuming. It is important to note that the VOLL for the "average curtailed customer" is not the same as the VOLL for the average customer across the market as a whole. For example, due to demand response and the existence of "priority lists", not all customers are equally likely to be interrupted due to a wholesale-level reliability event.

While efficient on average, setting the price at the average VOLL of involuntarily curtailed customers will not make every customer indifferent to service interruptions, because individual end users' VOLL span a wide range depending on customer class, income, industry, interruption duration, and a number of other factors that we discuss further below.

A. PJM interconnection and ISO New England

PJM and ISO-NE procure capacity to achieve reliability standards mandated by the North American Reliability Corporation (NERC).[99] This process includes three distinct steps:

1. Setting the reliability standard
2. Modelling the target reserve margin required to achieve the mandated reliability standard, and
3. Designing and implementing the capacity auction to procure the target capacity.

Each of these three steps requires modelling and engineering judgement calls, which tend to result in delivering more rather than less reliability. In the United States, customer outages due to supply shortages are rare, making up only 5% of major historical supply disturbances.[100]

In this section, we outline the process PJM and ISO-NE take to procure capacity and how implicit and explicit bias towards greater reliability affects their decision-making.

Markets for Electricity," *Wiley-IEEE Press*, 1, pp. 149-150; and Samuel Newell et al. (2012), ERCOT Investment Incentives and Resource Adequacy, Prepared for ERCOT, June 2012, p. 12.

99 NERC is a not-for-profit international regulatory authority operating in the United States, Canada and Baja California, Mexico. NERC is subject to oversight by the Federal Energy Regulatory Commission (FERC) in the United States and governmental authorities in Canada. NERC is governed by a Board of Trustee's which consists of independent members elected by NERC stakeholders and NERC's chief executive officer. See NERC (2019), *Governance*, and *About NERC*, accessed on 14 January 2019.

100 According to NERC, supply shortages made up 5.3% of disturbances that interrupted more than 300 MW or 50,000 customers, 1984 – 2006. See Paul Hines, Jay Apt, and Sarosh Talukdar (2008), "Trends in the History of Large Blackouts in the United States," *Proc. of the IEEE Power and Energy Society Generating Meeting*, 2008.

Step 1: Set Reliability Standard

Both PJM and ISO-NE procure capacity consistent with a 1-in-10 loss of load expectation (0.1 LOLE) standard mandated by NERC. This standard is interpreted as "one occurrence of outage in ten years".[101,102] The LOLE standard is a traditional engineering standard "rule of thumb" that has been used for many decades by utilities, system operators, and regulators across North America.[103]

The 1-in-10 LOLE standard does not explicitly account for either VOLL or other dimensions of customer attitudes towards the risk of outages, and studies have consistently demonstrated that the standard results in higher reserve margins than a standard whose reliability outcomes are aligned with VOLL.[104] As such, it could be considered implicitly loss averse. However, this loss aversion is not explicit on the part of NERC and is more an outcome of the level of reliability resulting from the engineering standard than any concrete theory of risk aversion.

For example, a Brattle study for the Federal Energy Regulatory Commission (FERC) found the 0.1 LOLE standard would result in a 15.2% reserve margin for a hypothetical RTO. In contrast, a 7.9% reserve margin would minimize societal costs, including VOLL costs, for a risk-neutral planner.[105] A 2018 Brattle report for ERCOT found that applying the 0.1 LOLE standard to Texas would require a reserve margin of 13.5%, as compared to an economically optimal reserve margin of 9.0%.[106] Studies have found that achieving a 0.1 LOLE standard implies a VOLL of US$200,000/MWh or higher, significantly above most empirical estimates of VOLL.[107]

---

101 PJM Interconnection (2015), *PJM Manual 20: PJM Resource Adequacy Analysis*, August 2015, Section 1.4.

102 ISO-NE (2015), *ISO New England Installed Capacity Requirement, Local Sourcing Requirements and Capacity Requirement Values for the System-Wide Capacity Demand Curve for the 2018/19 Capacity Commitment Period*, February 2015.

103 Johannes Pfeifenberger et al. (2013), *Resource Adequacy Requirements: Reliability and Economic Implications*, Prepared for FERC, September 2013.

104 The 0.1 LOLE standard also results in higher reserve margins than other common engineering reliability standards, such as 24 loss of load hours per 10 years (2.4 LOLH) and 0.001% expected unserved energy (0.001% EUE)

105 A 2.4 LOLH standard results in a reserve margin of 8.2% and a 0.001% Normalized EUE standard results in a reserve margin of 9.6%. See Johannes Pfeifenberger et al. (2013), *Resource Adequacy Requirements: Reliability and Economic Implications*, Prepared for the Federal Energy Regulatory Commission, September 2013, Table ES-1.

106 Newell et al. (2018), *Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region*, October 2018, p. 36.

107 Patton, D. (2013), "Comments of David B. Patton, PhD. Regarding State Policies Affecting Eastern RTOs." *FERC Docket No. AD17-11-000*, April 2017, p. 3. See also U.S. Federal Energy Regulatory Commission (2013), *Transcript of Technical Conference on Centralized Capacity Markets in Regional Transmission Organizations and Independent System Operators*, September 2013, p. 57.

Step 2: Modelling Target Reserve Margin Required to Achieve Reliability Standard

PJM and ISO-NE conduct reliability modelling to identify the level of capacity and reserve margin required to achieve the mandated 0.1 LOLE standard. This type of reliability modelling is common practice amongst most system operators worldwide. However, the traditional modelling techniques used in these types of studies have some limitations and necessary simplifying assumptions. Many of these limitations lead to results that tend to overstate the amount of capacity needed to achieve the reliability standard. One such example is that reliability modelling may not take into account all actions that operators can take to reduce the likelihood of outages, such as reducing load by ordering distribution networks to implement voltage reductions.[108]

On the other hand, reliability modelling limitations can also tend to understate or miss some reliability and security risks. For example, traditional reliability modelling is focused on procuring sufficient capacity to meet system peak loads, and does not take into consideration the nature of outage drivers that are not directly related to meeting peak load, such as flexibility or ramping needs. Traditional reliability modelling also may understate risks by not taking into account weather-driven and common-model failure risks amongst generators that can lead to "tail risks". These limitations were recently illuminated during an unexpected period of extreme cold in January 2014, in which 22% of generating capacity was simultaneously offline and demand for electricity was 25% higher than typical January peaks.[109] The abnormal level of supplier outages and customer load were both driven by abnormally cold temperatures. Apparent "over-procurement" of capacity is likely the result of these known modelling limitations (as well as concerns about unknown risks not captured by the model).

Step 3: Capacity Procurements via Centralized Capacity Auction

Both PJM and ISO-NE procure capacity through a centrally administered capacity auction. These auctions are nominally designed to achieve the target reserve margin on average across multiple years. By their nature, capacity auctions are highly administrative and require subjective decision making by the system operator that can result in the auction consistently over-procuring capacity relative to the target reserve margin required to achieve the 0.1 LOLE standard. These over-procurements relative to the target reserve margin result in further misalignment between reliability outcomes and VOLL.

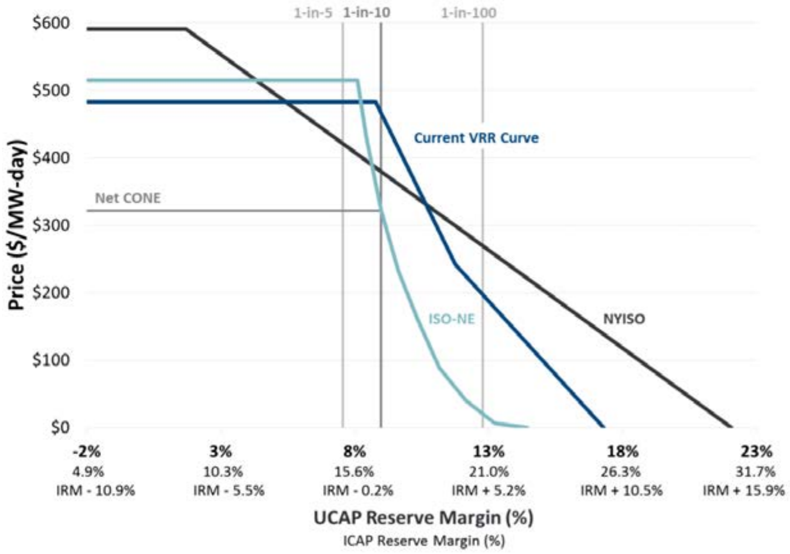
The capacity market demand curve is a key administratively-determined parameter that affects the expected level of capacity procured by the auction. Figure 3 illustrates the demand curves used by PJM, NYISO, and ISO-NE.

---

108 Johannes Pfeifenberger et al. (2013), *Resource Adequacy Requirements: Reliability and Economic Implications*, Prepared for the Federal Energy Regulatory Commission, September 2013, p. 9.

109 PJM Interconnection (2014), *Analysis of Operational Events and Market Impacts during the January 2014 Cold Weather Events*, May 2014.

Figure 2: Capacity Market Demand Curves in PJM, NYISO, and ISO-NE



Sources and Notes: Newell et al. (2018), *Fourth Review of PJM's Variable Resource Requirements Curve*, Prepared for PJM Interconnection, April 2018.

Current VRR Curve refers to the PJM capacity market demand curve.

ICAP: Installed Capacity

UCAP: Unforced Capacity (adjusted for outages and derates)

IRM: Installed Reserve Margin

Net CONE: Net Cost of New Entry

PJM Demand Curve

PJM's capacity demand curve has consistently resulted in over-procuring capacity relative the target reserve margin. Brattle's 2018 review found the current PJM demand curve would result in long-run reserve margins that exceed the target Installed Reserve Margin by 4.3%.[110] This over-procurement would in turn result in an average LOLE of 0.011 events per year, approximately an order of magnitude fewer outage events than the 0.1 LOLE target.[111] PJM has recently proposed adjusting the demand curve to result in average LOLE of 0.02 events per year and slightly better align reliability outcomes with the 0.1 LOLE standard.[112]

Several aspects of the demand curve shape including the price cap, estimated Net Cost of New Entry (Net CONE), forecast load growth, and slope result in the curve consistently over-procuring capacity.[113] Specifically, PJM intentionally designed the demand curve procure a 1% higher reserve margin on average than the target reserve margin. This 1% adjustment was made

---

110 PJM's target IRM for the 2021/22 base residual auction was 15.8% on an installed capacity (ICAP) basis, i.e. not derating generator availability for forced outages. See PJM Interconnection (2018), *2021/2022 RPM Base Residual Auction Planning Period Parameters*, 2018.

111 See Newell et al. (2018), *Fourth Review of PJM's Variable Resource Requirements Curve*, Prepared for PJM Interconnection, April 2018, pp. ix-x.

112 See Newell et al. (2018), *Fourth Review of PJM's Variable Resource Requirements Curve*, Prepared for PJM Interconnection, April 2018, pp. ix-x.

113 Net CONE represents a new plant's annualized fixed costs *net* of expected revenues from the energy and ancillary services markets. Rational new entry assumes that supply enters or exits the capacity market infra-marginally until the long-term average price equals Net CONE.

due to concerns that reliability modelling did not capture acute short-term supply uncertainty driven by regulatory-driven retirements, low gas prices, and the potential for Federal carbon policy. PJM argued these pressures "increased the likelihood of low reliability outcomes in the short term".[114] PJM's proposed 2018 FERC filing would remove the 1% adjustment as these short-term supply concerns have abated.[115]

ISO-NE Capacity Demand Curve

ISO-NE's capacity demand curve is designed to reflect the marginal reliability impact (MRI) of capacity.[116] The convex shape of the curve, as illustrated in Figure 4, reflects the expected improvement in reliability associated with adding incremental capacity.[117] The resulting curves are steeply sloped at lower capacity levels, reflecting larger improvements in reliability due to adding capacity, and less steeply slowed at higher capacity levels. The MRI curve is then developed into a priced demand curve by scaling the MRI-based curve such that it will achieve a price of Net CONE at the target reserve margin, and thereby attract the target level of capacity on average across years. The underlying design principal of the ISO-NE's MRI-based demand curves is that they allow all capacity procurements to be evaluated based on their marginal contribution to improving system reliability and bid cost.

Capacity Performance

Both PJM and ISO-NE have established additional so-called "Capacity Performance" incentives for capacity resources to be online in the event of an emergency that result in total supplier incentives in line with VOLL. Under Capacity Performance, capacity resources that under-perform relative to their awarded capacity amount (MW) during an emergency event are penalized and those that over-perform receive bonus payments. The $/MWh performance penalty rate (PPR) is calculated by PJM for each delivery year by dividing Net CONE by 30 hours, which is the number of emergency hours PJM anticipates the market will face at the target reserve margin. Capacity Performance mechanisms result in performance penalty rates of $3,000 to $5,500/MWh, which in conjunction with high energy prices during scarcity events results in total supplier incentives in line with estimates of VOLL.

ISO-NE Fuel Security

One of ISO-NE's largest reliability threats is increasingly winter reliability, specifically the availability of natural gas supply. However, ISO-NE's reliability mechanism and capacity product are designed to procure sufficient capacity to reliability achieve *summer* peak load. The limited scope of this reliability mechanism drove the need for out-of-market interventions when the Mystic plant and LNG terminal announced their retirement. The ISO believed this retirement would have caused significant winter reliability challenges in the Boston area.

---

114 PJM (2018), "Answer of PJM Interconnection L.L.C. to Protest and Comments", *FERC Docket ER-105-000*, p. 2.

115 PJM (2018), "Answer of PJM Interconnection L.L.C. to Protest and Comments", *FERC Docket ER-105-000*.

116 U.S. Federal Energy Regulatory Commission (2016), "Order Accepting Filing", *Docket ER16-1434-000*.

117 ISO New England (2016), *ISO New England Inc. and New England Power Pool Participants Committee, Docket No. ER16-000, Demand Curve Design Improvements*, Submitted to FERC.

ISO-NE's intervention may have many negative effects on the market, such as undercutting the profitability of other resources needed to achieve winter reliability. ISO-NE's intervention is currently under review at FERC and ISO-NE is developing a market-based solution. One potential market-based solution would be to define security and ancillary service products that incentivize winter operations to maintain fuel inventory and other winter supply, alongside forward investment products similar to capacity but focused on winter reliability needs.[118] Clear product definition will encourage low-cost, innovative solutions to achieve the need and not undermine market signals.

Takeaways for the NEM

- PJM and ISO-NE procure capacity in order to meet a 1-in-10 LOLE reliability standard. This traditional engineering standard does not consider VCR or outage preferences.
- PJM and ISO-NE carry higher reserve margins than a "risk neutral" energy-only market design whose prices reflect VCR, due to conservatisms built into the 1-in-10 LOLE standard, the design of the capacity auction, and reliability modelling techniques. Studies have found that achieving a 0.1 LOLE standard implies a VCR of US$200,000/MWh or higher, significantly above most empirical estimates of VCR.
- The winter of 2014 exposed reliability "tail risks" missed by traditional reliability modelling approaches (e.g. simultaneous high load and outages of many generators). PJM's conservatively high reserve margin allowed the RTO to ride through this extreme event without outages.

B. Electric Reliability Council of Texas (ERCOT)

ERCOT is an energy-only market and, unlike the rest of the United States, does not have a mandated reliability standard or reserve margin requirement.[119] Although ERCOT's energy-only design results in reliability outcomes generally aligned with customer risk preferences, ERCOT does not explicitly consider VOLL in their market design.

ERCOT has considered implementing a capacity market to ensure reliability, but ultimately has decided to keep the energy-only market design for two reasons. First, the energy-only design was viewed as more economical for customers as it attracts a level of reliability consistent with customer reliability preferences, as expressed as VOLL. Second, the energy-only design is viewed as a more targeted way to attract the fast-responding and flexible

---

118 Newell et al. (2018). *Market-Based Mechanisms for Winter Energy Security in NE*, Prepared for NextEra Energy Resources, December 2018.

119 Newell et al. (2018), *Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region*, Prepared for ERCOT, October 2018.

resources needed to resolve ERCOT's reliability and security concerns, which are increasing with growth in wind and solar. In contrast, a capacity auction that attracts capacity based solely on peak load requirements was viewed as a blunt instrument that may not attract the right types of capacity. In parallel, ERCOT has pursued a variety of other ancillary service and pricing reforms to ensure efficient pricing.[120]

ERCOT's Operating Reserve Demand Curve (ORDC)

Reserve margins and reliability outcomes in ERCOT are ultimately determined by suppliers' costs and willingness to invest based on energy and ancillary service market prices. These prices are set by market fundamentals and by the administratively determined Operating Reserve Demand Curve (ORDC) during tight market conditions. In pricing intervals where reserves are plentiful, the market price will reflect supplier offers (some of which must reflect marginal costs because of market power mitigation rules). When reserves are not plentiful, the ORDC adds an administratively-determined amount to the price to reflect the risk of reserves being exhausted. This approach creates a supply response to changes in energy market prices. Low reserve margins cause high energy and ancillary service prices and attract investment in new resources. Investment will continue until high reserve margins result in energy and ancillary prices too low to support further investment.[121]

The ORDC is an administrative price adder on energy and ancillary service markets that increases prices as supply conditions tighten and the likelihood of a supply shortage rises. Specifically, the administrative ORDC curves are calculated based on a loss of load probability at various levels of operating reserves, multiplied by an assumed VOLL of $9,000/MWh.[122] The $9,000/MWh parameter was set by the Public Utilities Commission of Texas (PUCT),[123,124,125] and is consistent with empirical estimates of VOLL .[126]

---

120 ERCOT (2019), *Future Ancillary Services Team*, accessed 15 January 2019.

121 Newell et al. (2018), *Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region*, October 2018.

122 ERCOT (2014), *ORDC Workshop*, April 2014, p. 17.

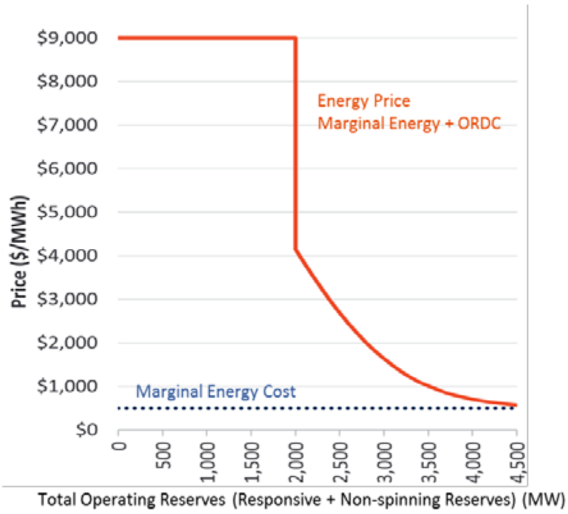
123 The PUCT is the regulator for the state of Texas' electricity, telecommunication, and water and sewer utilities. It implements legislation in the respective sectors, and offers assistance for resolving consumer complaints. The PUCT is a state agency, with the PUCT commissioners appointed by the governor of Texas. See PUCT (2019), *About the PUCT*, accessed on 14 January 2019.

124 PUCT (2012), *Electric Substantive Rules: Chapter 25, Subchapter S*, 2012. See also ERCOT (2017), *Study Process and Methodology Manual: Estimating EORM and MERM*, November 2017.

125 Prices can rise above $9,000/MWh in some circumstances due to local constraints. See ERCOT (2014), *ORDC Workshop*, April 2014, p. 9.

126 London Economics (2013), *Estimating the value of lost load*, June 2013.

Figure 3: Illustration of ERCOT Operating Reserve Demand Curve



Sources: Newell et al. (2018), *Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region*, October 2018.

Reliability Outcomes of ERCOT's Energy-Only Design

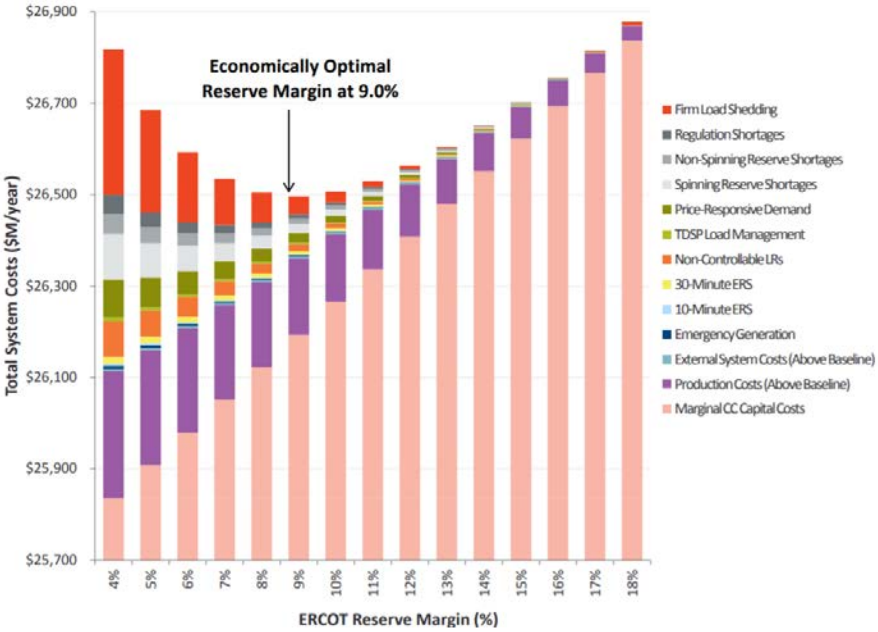
Through the implementation of the ORDC, the ERCOT market could be considered as generally aligned with customer risk preferences as it results in market prices that reflect both (1) rising loss of load probability as system security degrades, and (2) an offer cap generally aligned with estimates of VOLL.

A recent Brattle analysis commissioned by the regulator evaluated the performance of the ERCOT market design using a probabilistic reliability model. Performance, in terms of annual total system costs (including costs of customer outages valued at VOLL) and reliability outcomes were quantified across 57,000 simulations of different outcomes. Figure 4 shows the average simulated costs for each reserve margin as the red curve, which has a minimum total system cost at the risk-neutral reserve margin of about 9%. This would be the optimal level of reserve margin assuming risk neutrality.

The study found the market would attract an equilibrium reserve margin of 10.25%. This is 1.25% higher than the 'economically optimal' reserve margin of 9.0% that would minimize total societal costs, including costs associated with customer outages, based on the risk-neutral, probability-weighted result of the simulations. The study found total societal costs are relatively flat with respect to reserve margin near the minimum, with only modest variation between reserve margins of 7% and 11%, as illustrated in Figure 6. The design of the ORDC deliberately results in a higher level of reliability than would be justified in terms of minimizing expected economic costs, and a reserve margin in line with ERCOT's target reserve margin of 13.75%, set by the ERCOT Board of Directors.[127]

[127] ERCOT (2018), *Resource Adequacy*, 2018.

Figure 4: Illustration of ERCOT Operating Reserve Demand Curve

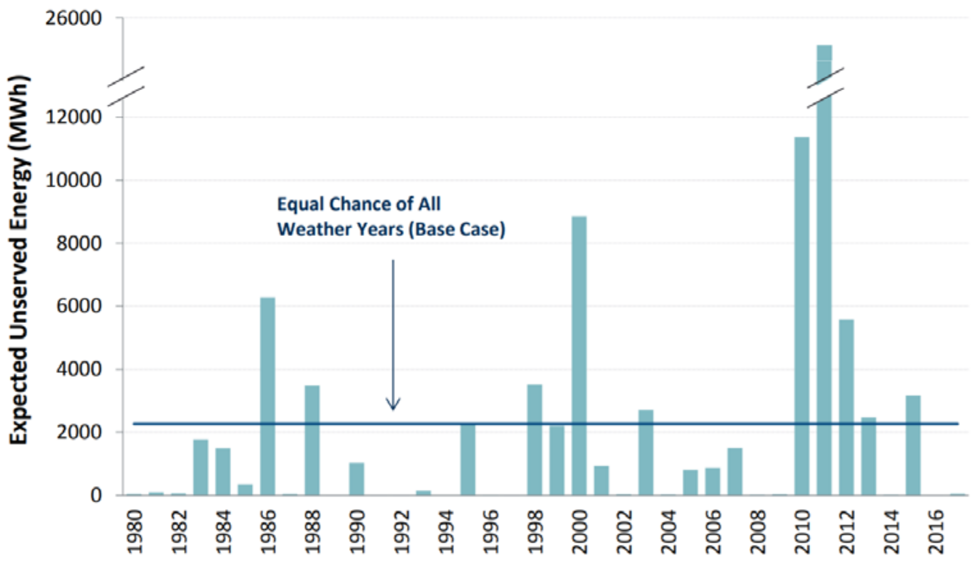


Sources: Newell et al. (2018), *Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region*, October 2018.

The study also compared the reserve margins resulting from ERCOT's energy-only market design to those reserve margins required to achieve either a 0.1 LOLE standard or a 2.4 Loss of Load Hours (LOLH) standard (where LOLE follows the same definition as in PJM and LOLH means the expected number of hours of interruption in a year, irrespective of magnitude) . The study found a 13.5% reserve margin would be needed to achieve a 0.1 LOLE standard, and a 9.2% reserve margin would be needed to achieve a 2.4 LOLH standard, as compared to the 10.25% market equilibrium reserve margin attracted by ERCOT's design.

The study also notes that average reliability and cost statistics assume risk neutrality on the part of customers and regulators, and can obscure the wide distribution of possible outcomes that may arise. Figure 7 illustrates how reliability outcomes vary with weather, as measured by expected unserved energy (MWh). Across the 38 weather years evaluated, 2011 weather conditions lead to 25,000 MWh of unserved energy, more than 10x higher than the equal-probability-weighted average of 2,300 MWh across all 38 years. This suggests that outage risk may be "fat tailed", with risk concentrated in high impact, low probability events.

Figure 5: Expected Unserved Energy by Weather Year at 10.25% Reserve Margin



Sources: Newell et al. (2018), *Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region, 2018 Update, Final Draft*, October 2018, Figure 11.

ERCOT's Emergency Reserve Service and Reliability Must Run Programs[128]

Although ERCOT nominally attracts capacity based on their energy-only market design and ORDC, ERCOT does have established programs to attract additional supply resources. These include the Emergency Response Service (ERS) program to attract demand-side resources, and reliability must run (RMR) programs for supply-side resources.

The ERS program procures additional demand-side resources beyond those attracted by the energy-only market design. A load-shedding event in the spring of 2006 prompted ERCOT to create the Emergency Interruptible Load Service (EILS) product, later renamed Emergency Response Service (ERS). During this event, ERCOT instructed transmission and distribution utilities to curtail load by 1,000 MW (2% of load) for nearly two hours.[129] The outage was driven by several confounding factors, including (1) extreme heat nearly 10°F above forecast, (2) 20% of the generation fleet out on planned outage, and (3) five major units tripping.[130] In response, the EILS program was proposed by ERCOT in September 2006.

ERS resources commit to be available for curtailment during times of scarcity via direct control of ERCOT. In exchange, ERS resources receive an availability payment, not unlike capacity payments in other jurisdictions. ERCOT procures ERS up to a $50 million annual expenditure limit. Expenditures are divided between each availability period in a year based on anticipated need for ERS.

128 For further details, see Spees et al. (2017), *Near-term Reliability Auctions in the NEM: Lessons from International Jurisdictions*, Prepared for AEMO, August 2017.

129 PUCT (2006), *Investigation into April 17, 2006 Rolling Blackouts in the Electric Reliability Council of Texas Region*, April 2006, p. 16.

130 ERCOT (2009), *ERCOT Emergency Interruptible Load Service*, July 2009.

ERCOT may hold a generator at risk of imminent retirement online through a reliability must run (RMR) contract. The RMR mechanism is primarily utilized to support system security, including voltage stability and other transmission system requirements, rather than short-term resource adequacy/reliability concerns.[131] ERCOT's nodal energy-only market design is intended to provide investment signals for suppliers to build capacity when and where it is needed to relieve scarcity conditions. RMR contracts are used rarely and only as short-term solutions; ERCOT must identify long-term transmission projects to address any concerns that necessitate a short-term RMR contract. Nonetheless, the effect of RMR is to increase the amount of supply available to the system beyond the level that would have been obtained purely in response to market prices. Similar to ERS resources, the output of RMR-contracted units is mostly withheld from the energy market by requiring these units to offer at the cap of $9,000/MWh into the day ahead market (the day ahead market is not mandatory for other resources). However, if the RMR unit is needed to relieve a binding transmission constraint, ERCOT may mitigate its offer down to a level far below the offer cap.

Takeaways for the NEM

- ERCOT's energy-only market design is generally aligned with the preferences of a risk-neutral customer, with a market price cap consistent with an estimate of the value of lost load or VCR.
- The resulting reserve margins are lower than they would be if ERCOT was meeting the 1-in-10 LOLE standard adopted by PJM and ISO-NE.
- ERCOT's energy-only market design assumes customers are risk-neutral to outage duration and magnitude, despite evidence that outage risk may be concentrated in a few, high impact events.
- However, ERCOT's emergency response service (ERS) and reliability must run (RMR) programs provide additional reliability and system security beyond that attracted through the energy-only market design, and the ORDC is also designed to deliver additional resources.

C. Ofgem (Great Britain)

In 2010, Ofgem conducted a year-long study into whether the then-current arrangements in Great Britain were adequate for delivering secure and sustainable electricity (and gas) over the next 10-15 years. The study showed that supply would be secure through 2015, after which significant new investments would be needed to deliver electricity securely and sustainably.[132] As part of the resulting Energy Market Reform, the Government introduced a Capacity Market (first delivery in winter 2018/2019) to drive new investments to ensure that sufficient capacity is available to meet demand.

---

131 An RMR contract was used to support short-term resource adequacy in 2011. See ERCOT (2016), *Reliability-Must-Run Procedures*, 2016.

132 Ofgem (2010), *Project Discovery: Options for delivering secure and sustainable energy supplies*, February 2010.

Step 1: Set Reliability Standard

In setting the reliability standard, the Department of Energy & Climate Change (DECC) considered the trade-off between the cost of providing additional capacity (measured by the net Cost of New Entry, or net CONE), and the benefit of a lower risk of interruptions (measured by consumers' VOLL).[133]

$$Reliability\ standard = \frac{Cost\ of\ new\ entry}{Value\ of\ lost\ load}$$

In 2013, DECC set a reliability standard of no more than three hours of LOLE per year.[134] This was estimated using a cost of new entrant capacity of £47,177/MW/year and a VOLL of £16,940/MWh. Due to the level of uncertainty around the estimation, DECC thought it more appropriate to round up the reliability standard from 2 hours, 47 minutes and 6 seconds to three hours. [135]

Step 2: Modelling Target Reserve Margin Required to Achieve Reliability Standard

National Grid, the electricity system operator in Great Britain, has statutory responsibility to produce an annual report to the Government recommending a level of capacity to procure through the Capacity Market. To do so, it employs DECC's Dynamic Dispatch Model to forecast needed capacity to meet the reliability standard under a range of scenarios and sensitivities.[136]

In addition to a Base Case, four scenarios are considered, which are developed through National Grid's Future Energy Scenarios' stakeholder process. Finally, sensitivities (changing only one variable at a time) around the Base Case are also considered to "cover uncertainty" by considering ranges for peak demand, demand side response, plant availability, interconnection capacity, weather, and wind levels. [137,138] The four current scenarios are as follows:

- **Community renewables**: High adoption of electric vehicles (EVs) and deployment of heat pumps, improved appliance energy efficiency (EE), greater use of demand response.[139]
- **Steady progression**: Limited improvements in EE, little electrification of heat, significant adoption of EVs .[140]

---

133 DECC (2013), *Annex C: Reliability Standard Methodology*, 2013, p. 7.

134 DECC (2013), *Annex C: Reliability Standard Methodology*, 2013, p. 8.

135 DECC (2013), *Annex C: Reliability Standard Methodology*, 2013, p. 8.

136 DECC (2012), *DECC Dynamic Dispatch Model*, May 2012, p. 3.

137 National Grid (2018), *Electricity Capacity Report*, May 2018, p. 5.

138 Reviewers of National Grid's approach (see below) have suggested that sensitivities should be run around the other scenarios as well, but this suggestion has not been implemented.

139 National Grid (2018), *Electricity Capacity Report*, May 2018, p. 24.

140 National Grid (2018), *Electricity Capacity Report*, May 2018, p. 25.

– **Two degrees**: Smart technology to be used alongside greater demand side response to manage peak electricity demand, improved EE.[141]

– **Consumer evolution:** Moderate roll-out of smart charging to accommodate the achievement of the 2040 transport target for no further petrol or diesel cars, some improvements in EE, more focus on local generation.[142]

Figure 6: National Grid's four Future Energy Scenarios

Sources: National Grid (2018), *Scenario Framework document*, 2018.

National Grid has explicitly decided not to include a HILP sensitivity (black swan events like nuclear type faults, or extreme cold weather) in its two Electricity Capacity Reports (2017, 2018) so far. This is because it has investigated nuclear type faults previously and "concluded that they were low probability and historically had been rectified ahead of the following winter". [143] Modelling extreme cold weather would "involve changing more than one element which violates the principles behind the sensitivities of only including credible outcome by changing one variable" and in extreme weather circumstances, the system operator would have recourse to the latent capacity on the system.[144] For these reasons and the assumption that "[policy] uncertainty around coal will be addressed through the non-delivery sensitivities",

[141] National Grid (2018), *Electricity Capacity Report*, May 2018, p. 24.

[142] National Grid (2018), *Electricity Capacity Report*, May 2018, p. 26.

[143] National Grid (2018), *Electricity Capacity Report*, May 2018, p. 40.

[144] National Grid (2018), *Electricity Capacity Report*, May 2018, pp. 40-41.

National Grid has decided not to include any "black swan" event sensitivities.[145] Nonetheless, National Grid uses a Least Worst Regret approach to determine the capacity to procure across the scenarios and sensitivities, to avoid assigning subjective probabilities to each scenario/sensitivity.

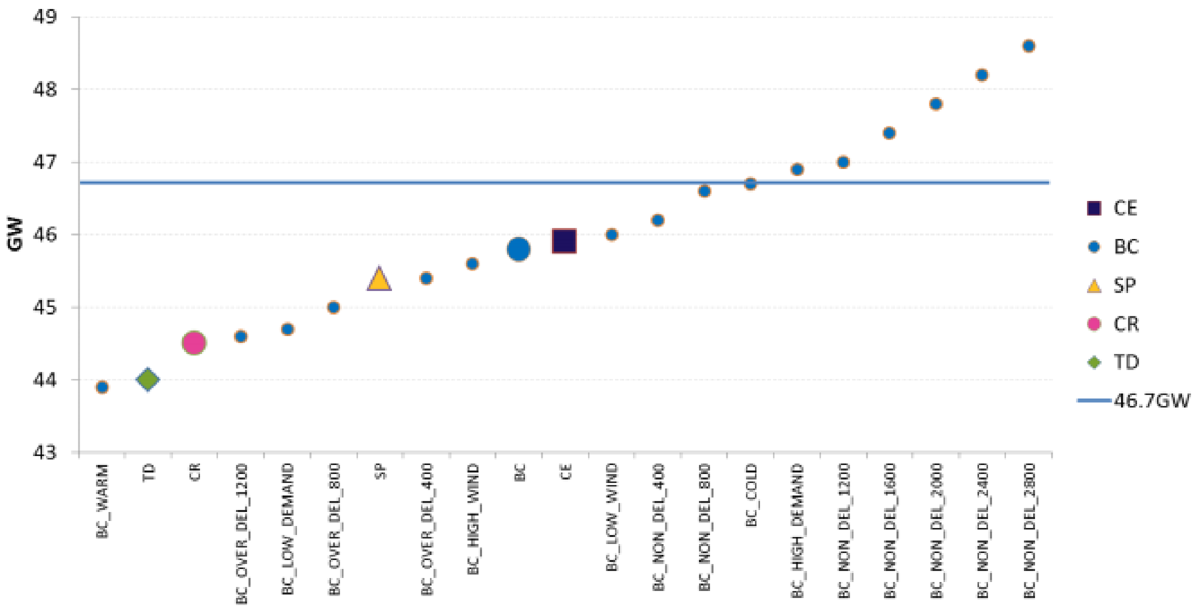
Step 3: Capacity Procurements via Centralized Capacity Auction

National Grid uses a Least Worst Regret approach to determine the capacity to procure across the scenarios and sensitivities it has modeled (where each case satisfies the reliability standard). This approach is a cost optimization of the potential outcomes that could occur if one scenario/sensitivity is assumed but another one occurs. It does so by assuming costs that are consistent with the reliability standard (a VOLL of £17,000/MWh and a net CONE of £49/kW/year).[146]

$$Total\ cost = Derated\ capacity\ secured * Net\ CONE + USE * VOLL$$

In the latest Electricity Capacity Report (2018), National Grid showed the range of de-rated capacity to secure in the auction for delivery in 2022/23 to meet the three-hour LOLE reliability standard across the scenarios and sensitivities. Individual scenarios are depicted in Figure 10 with larger markers/shapes. Sensitivities around the Base Case are identified using smaller dots.

Figure 7: Least Worst Regret recommended capacity to secure in 2022/23



Sources and Notes: National Grid (2018), *Electricity Capacity Report*, May 2018, Figure 24. CE = Consumer Evolution, BC = Base Case, SP = Steady Progression, CR = Community Renewables, TD = Two Degrees, capacity to procure = 46.7 GW.

The LWR approach yielded an outcome of 46.7 GW of capacity to procure (from the Cold Winter sensitivity). This means that if any scenario/sensitivity other than Cold Winter were to occur with this level of capacity having been procured, the maximum total cost taking into account the cost capacity and the cost of USE would be smaller than if the level of capacity had

145 National Grid (2018), *Electricity Capacity Report*, May 2018, pp. 40-41.

146 National Grid (2017), *Electricity Capacity Report*, May 2017, pp. 87-88.

been optimized on a different scenario. National Grid believes that the LWR approach is independent of the scenarios/sensitivities' likelihood of occurrence, and therefore "it can be used when the probabilities of these outcomes are unknown, providing that the cases considered cover a range of credible outcomes".[147]

An independent Panel of Technical Experts reviews National Grid's analysis every year to quality assure National Grid's modelling and approach.[148,149] In their 2016 report, the Panel critiqued National Grid's approach of essentially assigning equal weights to extreme sensitivities.[150] Instead, it proposed a probabilistic assessment, which assigned probabilities to each scenario/case. Using this approach, for the test year 2020/21, the Panel found that if the worst case were 33% likely to occur, the capacity to secure would be reduced by 0.9 GW relative to the amount procured under the LWR; similarly if the worst case were 50% likely to occur, the capacity to secure would be reduced by 0.6 GW.[151]

In response to the Panel's critique, National Grid commissioned academic consultants to review the current approach and to look at options for incorporating probabilities into the process.[152] Although this did not lead to a change in the methodology, National Grid acknowledged that Zachary and Wilson (2016) findings were important:

- **Bayesian decision-making not always possible:** Assigning probabilities to future conditions is not always possible because of insufficient data. While extreme weather conditions could be probabilistically quantified from historical data, significant non-delivery of contracted conventional generation is harder to quantify because the relevant data does not exist (would have to rely on expert judgment which is subjective)[153]
- **Boundaries by extreme sensitivities**: Zachary and Wilson have shown that the results of LWR approach is determined by the cost functions associated with the two most

---

147 National Grid (2017), *Electricity Capacity Report*, May 2017, p. 87.

148 EMR Panel of Technical Experts (2016), *Final Report on National Grid's Electricity Capacity Report*, June 2016.

149 The Panel of Technical Experts is appointed by government through an "open and transparent procurement process". The members consist of independent consultants with "diverse experiences of the electricity markets in Great Britain and Ireland, as well as knowledge of a wide range of generation technologies." See Department for Business (2017), *Energy & Industrial Strategy*, *Panel of Technical Experts*, July 2017, p. 8; GOV.UK, Panel of Technical Experts, accessed on 15 January 2019

150 "For example, whilst we accept the argument of National Grid's consultant statisticians Zachary and Wilson that it is reasonable to make extra allowance for extreme weather events, but that does not mean giving them 100% weight." See, EMR Panel of Technical Experts (2016), *Final Report on National Grid's Electricity Capacity Report*, June 2016, p. 46.

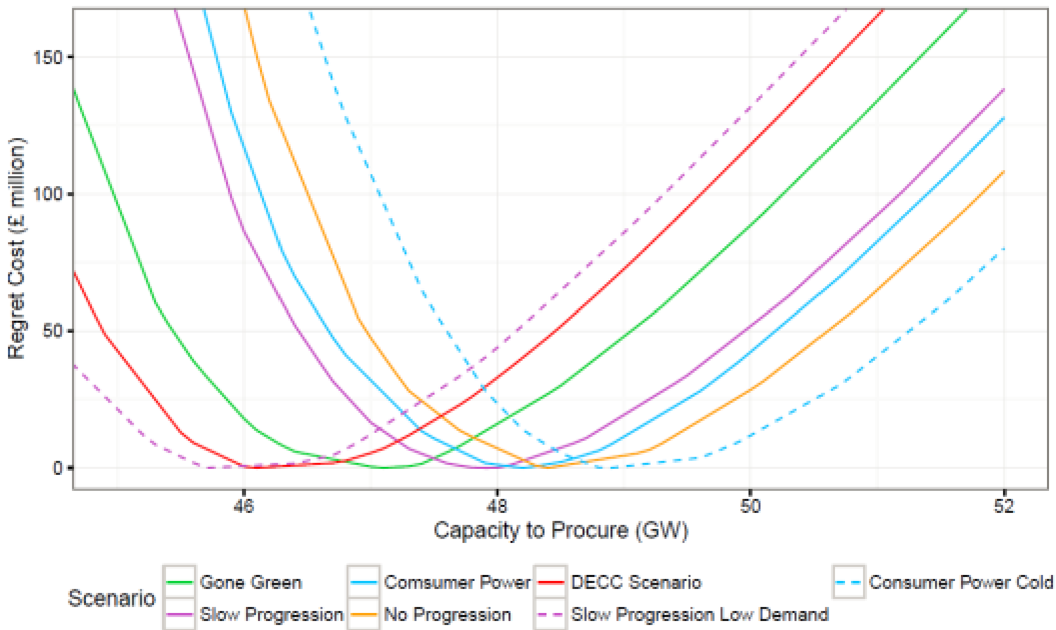
151 EMR Panel of Technical Experts (2016), *Final Report on National Grid's Electricity Capacity Report*, June 2016, pp. 43-44.

152 National Grid (2017), *Electricity Capacity Report*, May 2017, p. 88.

153 Stan Zachary (2016), *Least worst regret analysis for decision making under uncertainty, with applications to future energy scenarios*, August 2016, p. 2.

extreme scenarios (the most optimistic and most pessimistic ones), while being largely unaffected by other scenarios. The current approach to LWR yields the solution as the value corresponding to the point of intersection of the regret functions for the two extreme sensitivities (dotted lines in Figure 11)[154]

Figure 8: Regret cost functions



Sources: National Grid (2017), Electricity Capacity Report, May 2017, Figure 32.

– **Hybrid approach**: Zachary and Wilson believed (in line with the Panel's critique) that a fully probabilistic analysis would be preferable to the LWR approach; however, they also acknowledged that the assignment of probabilities would be wholly subjective. Thus, as a compromise between assigning probabilities to all scenarios/sensitivities and accounting for extreme events, they recommended a "hybrid" approach. Under this approach, probabilities can be assigned to extreme sensitivities (those with extreme capacity requirements) so that they do not exert undue influence.

As a result of the consultants' recommendation, National Grid performed an indicative hybrid approach analysis for 2020/21, and assumed probabilities of 0.011, 0.011, 0.009, and 0.009 for the four most extreme cases. It found only a difference of around 0.2 GW in the capacity to secure relative to the LWR approach, so decided not to implement the hybrid approach at this time.[155] In its 2017 report, the Panel acknowledged National Grid's trial run findings, though it did not think that it refuted its previous recommendation. Thus, the Panel again put forward its recommendation to apply "decreasing weights for more extreme sensitivities within the

[154] Stan Zachary (2016), Least worst regret analysis for decision making under uncertainty, with applications to future energy scenarios, August 2016, pp. 13-14.

[155] National Grid (2017), Electricity Capacity Report, May 2017, pp. 91-92.

LWR set of sensitivities".[156] This was not addressed in National Grid's latest (2018) Electricity Capacity Report.[157]

Auction results so far have "delivered low prices and more capacity has been procured". Because offers in the auction have generally been at prices lower than the assumed net CONE, more than the target quantity of capacity has been procured (the demand curve in the auction is sloping, permitting up to 1.5 GW more or less than the target amount to be procured). The net result of targeting capacity to meet a more demanding scenario than the Base Case, and procuring above the target, is a Base Case having less than two hours LOLE for the period covered by the auction.[158] Great Britain, thus like many other electricity markets, ends up delivering more capacity than required by the reliability standard.

Takeaways for the NEM

- A capacity mechanism was introduced in GB to replace the prior energy-only design. The capacity mechanism is designed to meet a reliability standard of 3 hours per year LOLE, which is in line with the level at which the costs of reserve capacity are about equal to the cost of interruptions.
- Modelling is done to derive a target quantity of capacity that would be consistent with the reliability standard. This modelling is not done purely in terms of expected USE or expected hours of interruption, but instead incorporates a "least worst regret" approach. The least worst regret approach is scenario based but each scenario is given equal weight rather than being weighed by probabilities.
- The least worst regret approach has been criticised for effectively being driven by "extreme" scenarios (ie, the most and least challenging scenarios), and for resulting in the procurement of more capacity than a risk-neutral probability-weighted approach.
- Procuring more reserves than suggested by a probability-weighted approach may implicitly correct for HILP events that have intentionally not been modelled.

---

156 EMR Panel of Technical Experts (2017), *Final Report on National Grid's Electricity Capacity Report*, July 2017, pp. 46-47.

157 National Grid (2018), *Electricity Capacity Report*, May 2018, p. 85.

158 National Grid (2018), *Modelling methods, Version 2.0*, July 2018, p. 23.

V. Implications for the NEM

In the NEM the primary mechanism for delivering wholesale-level reliability is the price mechanism: in response to high prices, there is investment in additional supply, and loads will consider reducing consumption at peak times. The NEM also has a backstop mechanism – the RERT – which can be used to procure additional supply (including demand response) if the price mechanism is not effective at delivering wholesale-level reliability. This could happen, for example, following an unexpected generator retirement and long lead times for investing in new supply. The long-notice and medium-notice RERT can only be used if AEMO determines that the NEM's reliability standard of 0.002% expected USE will not be met.[159] When AEMO forecasts expected USE and compares its forecast with the reliability standard, it does so by modelling a range of possible future outcomes and weighting the amount of USE in each outcome by an estimate of the probability associated with that outcome.

There are many settings in which individuals have to take decisions in relation to uncertain outcomes, such as purchasing insurance. Individuals' choices are often not consistent with a framework in which outcomes are weighted by the corresponding probability. Most electricity consumers in the NEM do not pay the spot price for electricity in the wholesale market, but rather buy electricity from retailers at a fixed price. This behaviour is risk averse.

In this report we have examined concepts of risk aversion and loss aversion to determine whether they might suggest an approach for triggering the NEM's backstop reliability mechanism different from comparing expected USE to the reliability standard, and specifically whether a different approach might be needed for managing risks associated with HILP events. We have examined underlying economic theories and looked at whether these concepts are reflected in decisions about wholesale-level reliability in other jurisdictions, or in decisions in other sectors of the economy.

A. Consumer risk preferences

In many circumstances, consumers prefer a certain outcome to an uncertain outcome, even if the latter is *expected* to be more favourable (where expected means a probability-weighted average). There is a large literature explaining how and why consumers make choices in which they display risk aversion, including the theory of expected utility with diminishing marginal utility of wealth, and prospect theory and loss aversion. Each of these may have implications for wholesale-level reliability, as we discuss below.

If consumers aim to maximize expected utility, and marginal utility of wealth is declining, they will be risk averse when faced with risks that are significant relative to total wealth. For

---

159 For a short-notice RERT (where AEMO has less than seven days of notice), AEMO triggers the RERT based on LOR declarations (which operationalise the reliability standard). See AEMC (2018), *Enhancement to the Reliability and Emergency Reserve Trader, Options Paper*, October 2018, pp. 10, 25.

example, purchasing insurance on a large asset such as a house is consistent with expected utility theory. This theory also explains why investors demand a risk premium for investing in risky assets (shares) rather than riskless assets (government bonds). This theory does not explain the magnitude of the equity risk premium, and it does not explain why consumers also appear to be risk averse in relation to very small risks.

We understand that AEMO's modelling includes a range of possible outcomes that, if they were to happen, would result in USE. Most of the outcomes with USE have relatively small amounts of USE, but others are much larger, associated with HILP events. If these HILP events could cause a significant amount of harm to consumers (significant in relation to consumers' overall wealth/income), then maximizing consumer utility might imply putting a greater weight on the impacts caused by HILP events. However, we think that outages associated with a lack of wholesale-level reliability would not give rise to consumer impacts that are significant in relation to overall wealth/income. For example, such outages rotate and thus an individual consumer need not be off supply for an extended period of time. We therefore consider that the risk neutral approach of forecasting expected USE is consistent with expected utility theory explanations of risk aversion, in relation to wholesale-level reliability.

In contrast to expected utility theory, prospect theory and the theory of loss aversion suggest that consumers will be risk averse in relation to risks that are small relative to overall wealth/income. This aspect of consumer behaviour is regarded as being a preference for avoiding losses: consumers feel happier knowing that they are not exposed to the risk of a loss, even if it costs them some money to avoid this risk. For example, this may be why consumers purchase insurance against relatively small risks, such as breakdown of electrical appliances.

Whether particular outcomes are perceived as being losses or a gain, will depend on the consumer's reference point, which is often modelled as their expectation of the future.[160] If there are two potential outcomes of a gamble and a consumer perceives one outcome as a loss and the other as a gain (relative to their reference point), then the consumer would be willing to pay a premium (relative to the probability weighted value) to avoid the gamble.[161] However, if the consumer saw both outcomes as being losses (relative to their reference point), then their behaviour would depend on other assumptions of prospect theory. In particular, if the consumer had diminishing sensitivity to losses and gains (a common assumption), then they would prefer a gamble between two losses over a certain loss. This means that under loss aversion with diminishing sensitivity, consumers are *risk seeking* when choosing between losses.[162] Another implication of prospect theory is that consumers place a higher weight on low probability events in their decision making.[163]

---

160 Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), p. 179.

161 Nicholas Barberis (2013), "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), p. 175

162 Justin Sydnor (2010), "(Over)insuring Modest Risks," *American Economic Journal: Applied Economics*, 2(4), p. 195

163 Ted O'Donoghue and Jason Somerville (2018), "Modelling Risk Aversion in Economics," *Journal of Economic Perspectives*, 32(2), pp. 100-102.

Predicting how a loss averse customer would assess a change in the likelihood of a HILP wholesale-level reliability event is complex and will depend on both the severity of the event and the customer's reference point. First, although customers may over-weight low probability risks, more than 99% of outages actually experienced by customers is for reasons other than wholesale-level reliability.[164] This implies that if customers do not experience discernible differences between HILP reliability events and other outages, they may not see them as low probability risks and may therefore not be particularly concerned about them. Unless there are other unforeseen complications, such as a power system security event, wholesale-level reliability events are managed through rotational load shedding. Thus, even if an outage is prolonged, the impacts are spread across many consumers, with no individual experiencing a sustained outage.

Second, consumer assessments of the risk of HILP events would depend on their reference point. If consumers perceived the historical occurrence of USE events under the current framework to be their reference point, considering a new approach with additional reserves would involve choosing between two losses: an increased risk of HILP events, or paying for additional reserves to reduce uncertainty. In this case, loss aversion with diminishing sensitivity would predict that they would prefer to take the risk of increased HILP events over paying for the additional reserves with certainty. If however, the consumer expectation is that some mechanism would be employed to eliminate any increased probability of HILP events, then their reference point would be different. In this case, any reduction in the availability of reserves, relative to expectations, would be a gain (lower payments), and any increase in the risk of a HILP event, relative to the expected approach, would be a loss. If this were the case, then consumers would want a level of RERT that met their initial expectations.[165]

It is possible that the average consumer would prefer to buy insurance against the risk of outages that could be caused by a lack of wholesale-level reliability. There is no straightforward way for consumers to express their preferences. For example, so far as we know, there are no insurance products which would compensate customers who are interrupted during a load shed event caused by wholesale-level reliability,[166] In order to discover consumer preferences over such risks, therefore, it would be necessary to conduct survey work.

---

[164] Between 2007-08 and 2016-2017, 0.23% of supply interruptions were reliability-related. See AEMC (2018), *Reliability Frameworks Review, Final Report*, July 2018, Figure 2.1.

[165] This is discussed in the context of insurance premiums by Schmidt (2016) and Sydnor (2010). Both authors use prospect theory to explain why consumers prefer low deductibles, or full insurance when possible, when the extra cost of reducing the deductible is relatively high and the gain, relative to the cost of the insured asset is relatively low. See Ulrich Schmidt (2016), "Insurance Demand under Prospect Theory: A Graphical Analysis," *The Journal of Risk and Insurance*, 83(1), pp. 77-89 and Justin Sydnor (2010), "(Over)insuring Modest Risks", *American Economic Journal: Applied Economics*, 2(4), pp.177-179.

[166] The fact that insurance products do not exist for wholesale reliability events does not in itself indicate that consumers would not wish to purchase such a product if it existed, particularly if the risk of wholesale reliability events were to increase relative to the status quo. Since insurance markets do not exist, consumers may not think about asking for insurance as their first response to

It is worth noting that mechanisms other than procuring additional reserves could reduce consumers' risks in relation to wholesale-level reliability. For example, if consumers were compensated for USE to the point where they were indifferent, there would be no risk.[167] Theoretically, the amount of compensation required would be the marginal VCR. If the MPC equalled the marginal VCR, then a consumer, paying the wholesale price would be indifferent between a wholesale-level reliability event resulting in load shedding and paying for the energy at the MPC. Of course, the average consumer procures electricity from a retailer at price that is far below the MPC. This results in the consumer losing the marginal VCR (less their much smaller retail price), while the retailer, who no longer has to purchase power at the MPC, gains by the MPC less the retail price. If the retailer refunded the MPC to the customer (and the MPC were equal to the marginal VCR), then there would be no risk to customers from wholesale-level reliability events.

B. Risk aversion and risks from HILP events in other settings

We have looked at how risk aversion is incorporated into the management of risks from HILP events in settings other than electricity reliability, specifically in relation to investment in flood defence, analysing risks associated with climate change, and insurance markets (see Appendices A and B for climate risk analysis and flood defence, respectively). We found that in many contexts assessments seem to be made on the basis of expected costs and benefits – ie, without incorporating risk aversion or allocating special weight to the risks of HILP events. However, sometimes additional weight may be put on the possible impacts of HILP events, either because of risk aversion or because of concerns that models may not adequately capture the impacts or probabilities of HILP events.

Flood defences are often planned on the basis of expected benefits (ie, avoided flood impacts). We did find that researchers in the Netherlands had considered incorporating risk aversion (utility maximization) into their models – however, they decided not to do so because government-provided compensation is available to those who suffer flood damage. If such compensation were not available, the researchers might have included a risk premium in their analysis, which presumably would have resulted in stronger flood defences. This is consistent with the discussion above, where utility maximization requires a risk premium for risks that are significant in relation total wealth (such as might be seen in flood damage to a home).

Policies designed to reduce or mitigate the impacts of climate change need to deal with great uncertainty, long time horizons, and HILP events (some of which could have catastrophic impacts, such as large changes in sea level). Nonetheless, it is common for policy analysts to proceed on the basis of expected costs and benefits. For example, commonly-used values of the social cost of carbon (SCC) are based on models of expected costs that weight HILP by estimated

dissatisfaction with increasing outage risk. Instead, they may signal their preference for reducing risk in other ways, such as through increased complaints about perceived outage risk.

167 This is similar to the Load Shedding Compensation Mechanism discussed by the AEMC in relation to compensation for load shedding due to reliability reasons. See AEMC (2018), Wholesale Demand Response Mechanism, Consultation Paper, November 2018, Appendix D.

probabilities. However, analysts also provide sensitivities which put more weight on HILP events. For example, values of the SCC are available that represent the less likely, but still possible impacts that would occur at the tail of the distribution of possible outcomes (95th centile), rather than just the average. Policymakers are also sometimes encouraged to focus on this type of "tail risk" – ie, reducing the probability of catastrophic impacts.

Demand for insurance products demonstrates the importance of risk aversion as a driver of human behaviour and is consistent with the underlying theories explaining why risk aversion arises. Consumers often buy insurance against risks that have relatively low impact but reasonably high probability, such as warranties on white goods. This is consistent with loss aversion. Consumers also buy insurance against high impact low probability events. For example, the fact that adults with dependents tend to hold life insurance is consistent with maximizing expected family utility, but not consistent with maximizing expected family wealth. Consumers tend not to buy insurance against types of catastrophic events with low probabilities that are difficult to estimate, such as earthquake or flood damage. This is inconsistent with theories of risk aversion, but could be due to confounding factors such as the cost of acquiring the information needed to assess such risks.

We think that HILP wholesale-level reliability events should not have significant impacts on consumers because of the design of rotating outages used to manage such events. However, it is also possible that conditions leading to wholesale-level reliability events could also result in elevated risks of interruptions for security reasons, which can be more widespread and longer-lasting. There may also be considerable uncertainty in models of the likelihood (and impacts) of these outages. In other settings (eg, climate risk analysis), these factors can be addressed by putting greater weight on HILP when analysing risk management options.

C. Reliability frameworks in other jurisdictions

We have examined the reliability standards in other jurisdictions, and the way in which those standards are implemented in practice. We looked at three jurisdictions with capacity mechanisms (PJM, ISO-NE and Great Britain) and one energy-only jurisdiction (ERCOT). Two of the four jurisdictions have reliability standards which are more stringent than an "efficient" standard that equates the cost of increasing wholesale-level reliability with the value to consumers of making that increase. In addition, in all four jurisdictions the reliability frameworks ultimately result in procuring more resources than system modelling shows is needed to meet the reliability standard.

- PJM and ISO-NE both aim to meet a standard of "1 in 10" loss of load expectation. They interpret this to mean one wholesale-level reliability outage per ten years, or an expected probability of having an outage of 0.1 per year (0.1 LOLE). This standard is significantly higher than one that equates the cost of improving reliability with the benefits to consumers of doing so. For example, it has been estimated that this 0.1 LOLE standard implies a VCR of about USD $200,000/MWh.
- In contrast, both ERCOT and Great Britain aim to meet an "efficient" standard consistent with a much more reasonable VCR. In Great Britain the standard is 3 hours per year LOLE, and ERCOT has an "economically optimal reserve margin", consistent with VCR.

- Three of the four jurisdictions have designs which mean that additional resources are available. For example:
    - In PJM the demand curve in the capacity mechanism is "right shifted", and the amount of capacity procured is estimated to be consistent with a standard of 0.01 LOLE
    - In ERCOT the system operator spends $50m per year on an emergency demand response mechanism, which is capable of delivering additional resource (if the market price cap is reached), right-shifts the scarcity pricing curve (ORDC), and can also offer generators reliability must run contracts to delay retirements
    - In Great Britain, the system operator models the amount of capacity needed to meet the 3 hrs LOLE standard, but the modelling is scenario based and the scenarios are not probability-weighted. This results in additional capacity being procured.

To a certain extent, apparent "over-procurement" may be the result of concerns that reliability modelling does not adequately capture all of the risks facing the system, or that some of the risks associated with HILP events cannot be reliably estimated. For example, PJM's system is principally designed to meet a summer peak demand, but a cold snap in 2014 resulted in high demand that coincided with large amounts of generator outage. An alternative explanation would be that there is a principal–agent problem: principal–agent problems arise where one entity acts on behalf of another, but incentives of the principal and agent are not perfectly aligned. In this case, institutions are acting on behalf of customers in relation to wholesale-level reliability, but these institutions may face incentives that elicit a risk averse response, irrespective of underlying consumer risk preferences. For example, system operators may face little upside from reducing system costs, but may have considerable downside from system outages in terms of reputational risk. A recent internal review of the Australian public sector found that there was dominant culture of risk aversion across various Australian government entities. [168] The internal review makes it clear that risk aversion comes from a culture that punishes losses and ignores gains.

Of the four jurisdictions we examined, one (Great Britain) has a modelling approach which explicitly rejects a "probability weighted" expected outage metric. The modelling in Great Britain uses a "least worst regret" approach, ostensibly because this approach allows a range of possible future scenarios to be incorporated without assigning probabilities to each one. Although the system operator's academic advisers have criticized this approach, for example because it is driven by the design of the more extreme scenarios, the system operator continues to use the least worst regret approach.

D. Conclusions

The current reliability framework in the NEM includes assessing *expected* unserved energy (USE) against a reliability standard. Expected USE consists of an estimate of USE in each of

---

168 Barbara Belcher (2015), *Independent Review of Whole-of Government Internal Regulation – Volume 1: Recommendations*, August 2015, p. 22.

many possible future outcomes, multiplied by the probability of that outcome occurring. For this report, we assume that, across the range of relevant outcomes, USE is a good measure of the relative impacts of the outages on customers. With this assumption, if risk aversion is relevant in the context of wholesale-level reliability, risk aversion would imply that consumers would be willing to pay for additional insurance to manage the risk of HILP events—such as additional reserves—in circumstances where the current reliability standard is met.

Maximising expected utility is different from maximising wealth in that it accommodates consumer preferences. A decision maker whose objective is to maximise expected utility should consider adopting a risk-averse approach if consumer preferences in the relevant context include risk aversion. This is sometimes seen in other contexts, such as climate change or flood defence, where HILP events can have very significant impacts. It is also consistent with consumer demand for certain types of insurance.

Different explanations for why consumers display risk aversion have different implications for the circumstances in which risk aversion might be relevant in a particular policy setting. Expected utility theory explains risk aversion as a function of the diminishing value of increased wealth. Generally, we would not expect this to be relevant in the context of wholesale-level reliability, unless the possible impacts of HILP events on consumers are large relative to consumers' total wealth. This seems unlikely (for example, in contrast to the possible impacts of flooding or fires). We have not seen any characterizations of the possible impacts on consumers of HILP wholesale-level reliability events, if they were to occur. However, we note that the system of rotating load shedding is designed to spread out these impacts so that individual customers should not be off supply for extended periods.

Unlike utility maximization, prospect theory and loss aversion can explain why consumers make risk-averse choices in relation to risks that are small relative to total wealth. For example, loss aversion explains why consumers purchase insurance against relatively low impact, high probability events (such as warranties on white goods).

Loss aversion could be relevant for assessing HILP wholesale-level reliability events if such events could give rise to amounts of USE that would be significant relative to what consumers might already anticipate from other sources (eg, distribution network faults). For example, if the circumstances in which wholesale-level reliability HILP events arise are also associated with increased system security risks, greater amounts of USE and greater customer impacts may be relevant. In this case, survey work could be undertaken to assess consumer preferences towards wholesale-level USE risks. This would be different from current surveys aimed to assess the value of customer reliability, but instead would focus on preferences over the risk of a USE event caused by wholesale-level reliability. For example, customers could be asked about their willingness to pay to reduce the risk of HILP wholesale-level reliability events, or the compensation they would be willing to accept if this risk increased. Since a loss averse customer's reference point is central to how they characterise the potential gains and losses from a risk, it would be extremely important to characterise accurately the underlying risk faced by the customer. If for example, HILP wholesale-level reliability events are indistinguishable from other outages from a customer perspective, then any changes in risk should be expressed relative to the baseline of total customer outages.

We have not seen risk aversion or loss aversion explicitly incorporated into reliability frameworks in other jurisdictions, nor have we seen any explicit discussion of consumer risk preferences. However, none of the four jurisdictions we examined uses an unweighted balancing of expected costs and benefits in its reliability framework. All four either target a reliability standard that is much higher than justified on the basis of the expected benefits of avoided outages, or procure more resources than required to meet the standard, or both. Two different factors are cited by system operators to explain (at least partly) why these jurisdictions approach wholesale-level reliability in this way. First, system operators recognize that the modelling underpinning the reliability framework does not adequately describe all of the possibly-relevant HILP wholesale-level reliability events (either it does not fully capture the impacts, if the event occurs, or it does not accurately estimate the probability of the event, or some possible HILP events are not modelled at all). Since the models do not capture all of these aspects, the system operator tends to procure additional resources. Second, system operators are often concerned about system security risks, and procure resources that can be used to manage both security risks and wholesale-level reliability risks. In such circumstances, we would recommend that security needs should be addressed directly, because the type of resource or service that can most efficiently address security needs may be different from that which can efficiently address wholesale-reliability needs.

A final possibility is that reliability frameworks may reflect a type of principal-agent problem, where institutions do not follow consumer risk preferences. This could be due to a perception that insufficient reliability would be costly for both institutions and customers, whereas procuring additional resources to increase reliability is costly for customers but not the institutions.

Appendix A: Examples of risk aversion from climate risk analysis

The field of climate risk analysis seeks to quantify the potential environmental, human health, and economic damages of greenhouse gas (GHG) emissions to inform policymakers and regulators allocating scarce resources between policy objectives of reducing climate emission and other objectives.

Climate risk analysis poses many challenges that parallel those of power system reliability planning, including:

- The need to understand a complex underlying system in the face of limited historical empirical evidence, requiring complex analysis, subjective modelling decisions and simplifying assumptions, and treatment of uncertainty
- The need to value damages created by unfavourable outcomes (increased global temperatures in climate analysis, customer outages in reliability planning), which requires subjective decisions on how to value damages and how to discount future damages

Researchers have begun to question if established climate risk analysis approaches appropriately accommodate risk-aversion and loss-aversion. Below we summarize standard approaches to climate risk analysis, recent developments, and how risk- and loss-aversion are considered.

Traditional Climate Risk Analysis Methods

The ultimate goal of climate risk analysis is to understand the social cost of carbon (SCC), or the external cost posed by unwanted climate change due to the production of CO_2 and other GHGs.[169] SCC is usually expressed in units of $/ton, reflecting the present value of future environmental damages created by emitting an additional ton of CO_2. By distilling complicated climate and economic analysis into a single metric, the SCC can be used by policymakers to value reductions in climate change and compare the effectiveness of reducing GHG emissions to other policy objectives.

At the highest level, estimating the SCC involves multiple steps:

- Simulating how GHG emissions today will affect the future climate;
- Translating changes in climate to consequences for human health, the environment, and the economy; and
- Valuating future damages and discounting damages appropriately.

[169] Robert Pindyck (2013), "Climate Change Policy: What do the Models Tell Us?," *National Bureau of Economic Research*, Working Paper 19244, July 2013.

Integrated Assessment models (IAMs) are the primary tool researchers use to develop SCC estimates. IAMs "integrate" climate science models of GHG emissions and their impact on temperatures with economic models of abatement costs and consequences of emissions for the future economy.[170] The models rely on reduced-form climate and economic models of varying complexity and provide a probabilistic estimate of SCC that accounts for uncertainty in key inputs and assumptions. IAMs allow users to consider risk aversion in multiple ways, including the choice of discount rate applied to future damages, and the choice of SCC from a distribution of outcomes.

Use of SCC by the U.S. Government

The U.S. Federal Government, under the administration of President Barack Obama, formed the Interagency Working Group on Social Cost of Greenhouse Gases (IAWG). The IAWG's purpose was to develop SCC estimates that would "allow agencies to incorporate the social benefits of reducing carbon dioxide emissions into cost-benefit analyses of regulatory actions".[171]

The IAWG did not recommend a single SCC value, but instead reported a range of values such that decision makers could account for their own risk-aversion preferences when valuing future climate damages. Specifically, the IAWG reported SCCs for three different discount rates: 2.5%, 3%, and 5%, as shown in Table 1. The IAWG reported both the expected (probability-weighted average) value of SCC under each discount rate, as well as the 95th percentile of the resulting probability distributions from the SCC estimates at a 3% discount rate. These values were derived from the underlying probability distributions resulting from each IAM, as shown in Figure 2. These distributions are "fat-tailed", meaning they imply an asymmetric risk of very damaging climate outcomes. IAWG acknowledges this asymmetric risk and its implications for risk-averse policymakers by reporting the 95th percentile results, stating:

> "there is extensive evidence in the scientific and economic literature on the potential for lower-probability, but higher-impact outcomes from climate change, which would be particularly harmful to society and thus relevant to the public and policymakers."[172]

---

170 Robert Pindyck (2013), "Climate Change Policy: What do the Models Tell Us?," *National Bureau of Economic Research*, Working Paper 19244, July 2013.

171 Interagency Working Group on the Social Cost of Greenhouse Gases, United States Government (2016), *Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866*, August 2016.

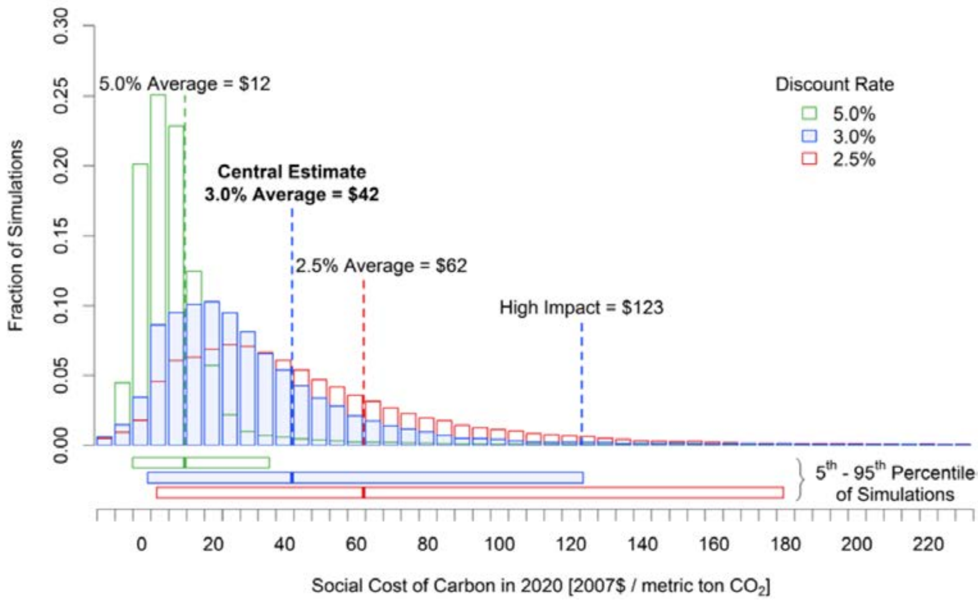
172 Interagency Working Group on the Social Cost of Greenhouse Gases, United States Government (2016), *Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866*, August 2016.

Table 1: Social Cost of CO2, 2010 – 2050 (in $US 2007 dollars per metric ton of CO_2)

Year	5% Average	**3% Average**	2.5% Average	High Impact (95th Pct at 3%)
2010	10	**31**	50	86
2015	11	**36**	56	105
2020	12	**42**	62	123
2025	14	**46**	68	138
2030	16	**50**	73	152
2035	18	**55**	78	168
2040	21	**60**	84	183
2045	23	**64**	89	197
2050	26	**69**	95	212

Sources: Interagency Working Group on the Social Cost of Greenhouse Gases, United States Government (2016), *Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866*, August 2016.

Figure 9: Frequency Distribution of SCC Estimates for 2020



Sources: Interagency Working Group on the Social Cost of Greenhouse Gases, United States Government (2016), *Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866*, August 2016.

Limitations of IAMs and the risk of catastrophic climate change

In recent years, academics and researchers have become concerned that IAMs may provide a flawed evaluation of future climate risk and false precision in a deeply uncertain problem. In particular, researchers have noted concern that IAMs understate the risk of catastrophic climate change that would represent an existential threat to the future global economy.[173] Catastrophic climate change outcomes are unlikely "tail events" in the range of potential future outcomes, but may represent the preponderance of risk to the global economy. IAMs are structurally limited in their ability to assess such tail risks that stretch the distribution of

173 Robert Pindyck (2012), "The Climate Policy Dilemma", *National Bureau of Economic Research*, Working Paper 18205, July 2012.

potential outcomes to the right, as they assume a relationship between emissions and loss of GDP that is based on conventional wisdom and as such may not reflect non-linearities that could result in much higher temperatures and loss of GDP.

In light of the potential threat of catastrophic climate change, some researchers have recommended climate risk analysts abandon integrated assessment modelling and instead focus on quantifying tail risks. Researchers have emphasized the need for "rough, subjective estimates" of the likelihood and consequences of catastrophic outcomes that are clearly presented as less precise than those SCC estimates from IAMs. Researchers could inform these estimates through eliciting the judgment of experts in the field.[174] Focusing on extreme outcomes would allow GHG abatement policy to be thought of as a form of insurance to guarantee low-probability catastrophes will not occur (or will be less likely).

Takeaways for the NEM

- As in reliability planning, climate risk analysis involves understanding a complex underlying system and valuing impacts from unfavourable outcomes
- The Social Cost of Carbon (SCC) quantifies the economic benefit of reducing GHG emissions. Policymakers often rely on an SCC that reflects the expected value of future climate outcomes, despite evidence that climate risk is "fat tailed"
- Researchers have recently suggested climate risk analysis should focus on the risk of catastrophic climate outcomes and climate mitigation be thought of as an insurance policy against such outcomes. This amounts to putting a greater weight on HILP events.

Appendix B: Examples of risk aversion from flood defence

Flooding is a major natural disaster affecting some 520 million people and causing $50 to $60 billion of damages every year.[175] The four main categories of flooding are:[176]

- **River flooding**: when a river cannot cope with the water draining into it from the surrounding land

[174] Robert Pindyck (2012), "The Use and Misuse of Models for Climate Policy", *Review of Environmental Economics and Policy*, 11(1), Winter 2017.

[175] Siegfried Demuth (UNESCO) (2013), *Preface to the International Hydrological Programme Book Series*, 2013, p. 1.

[176] Local Government Association (2019), *Managing Flood Risk*, accessed on 14 January 2019. See also Environment Agency (2009), *Flooding in England: A national assessment of flood risk*, 2009, p. 7.

- **Coastal flooding**: when low atmospheric pressure coincides with a high tide resulting in a tidal surge
- **Surface water flooding**: when heavy rainfall overwhelms the drainage capacity of the local area. This is difficult to predict and pinpoint, much more so than river or coastal flooding
- **Ground water flooding**: when water levels rise above the surface

The field of flood risk analysis seeks to quantify the potential economic, social, and environmental damages of flooding to inform national and local governments allocating scarce resources between the dual objectives of reducing flooding in the long-term and limiting the impacts of flooding when it occurs.

Flood risk analysis poses many challenges that parallel those of power system reliability planning, including:

- The need to predict impacts of an adverse event happening, requiring a modelling framework to be in place, an understanding of the geographic scope of the event, and a consistent treatment of uncertainty
- The need to value damages created by unfavourable outcomes, which requires subjective decisions on how to value damages and how to discount future damages

Flood risk management offers valuable insights into how expected costs and benefits translate into real-life investment decisions. There are two main types of approaches to flood risk management: (1) achieving a specified level of flood protection (much like the reliability standard); or (2) relying on analyses of risk to inform decision-makers about the relative cost-effectiveness of various options.[177] In this section, we examine approaches to flood risk analysis and investment in flood defenses, using The Netherlands (protection standard approach) and Great Britain (cost-effectiveness approach) as case studies.

177 U.S. Army Corps of Engineers (U.S.), Rijkswaterstaat (the Netherlands), Ministry of Land, Infrastructure, Transport and Tourism (Japan), Environment Agency (UK) (2011), *Flood Risk Management Approaches*, September 2011.

The Netherlands: 1 in 100,000 standard

The Netherlands is located in a low-lying delta formed by three major rivers. One-third of the country is below sea level, and two-thirds are vulnerable to flooding.[178] The Dutch flood defense mechanism is primarily consisted of dike ring areas (connected systems of embankments, dunes, and structures), for which the level of flood protection is specified by law.[179] Until recently, flood protection standards were defined as the frequency of exceedance in a year,[180] and these standards varied from 1:1,250 for the areas along the upper reaches of River Rhine and River Meuse to 1:10,000 for the most densely populated areas in the western part of the country.[181]

In 2017, the Netherlands changed its flood protection standard to a risk-based "life protection level of 1:100,000 years of becoming a flood casualty for every citizen living behind levees or dunes".[182] The new standard would translate to flood exceedance standards of between 1:2,000 per year (in the central area of the rivers Rhine and Meuse) and 1:10,000 per year, (in tidal river areas and in the central part of Holland).

The new risk-based flood protection standard was developed using a model that solves for the economic optimal dike (embankment) heights.[183] With climate change and economic growth (driving development in flood plains), flood probability (solid line in Figure 3) is likely to increase (though not in a monotonic manner). The model dynamically predicts and tracks the likelihood of flood risk and optimizes flood protection investments accordingly. That is, over time flood probability increases until a maximum "tolerable" flood probability is reached, at which point an investment is made, decreasing the probability back to the design standard. Due to fixed costs of installing flood protection measures, the optimal degree of reinforcement can be quite large, meaning that after the project is completed, the flood risk will have considerably reduced. After the investment, the flood probability increases again over time until the tolerable level is reached, and so on.

---

178 Chris Iovenko (Earth Magazine) (2018), *Dutch Masters: The Netherlands exports flood-control expertise*, October 2018.

179 JM Kind (2014), "Economically efficient flood protection standards for the Netherlands," *Flood Risk Management*, 7, pp. 103-117.

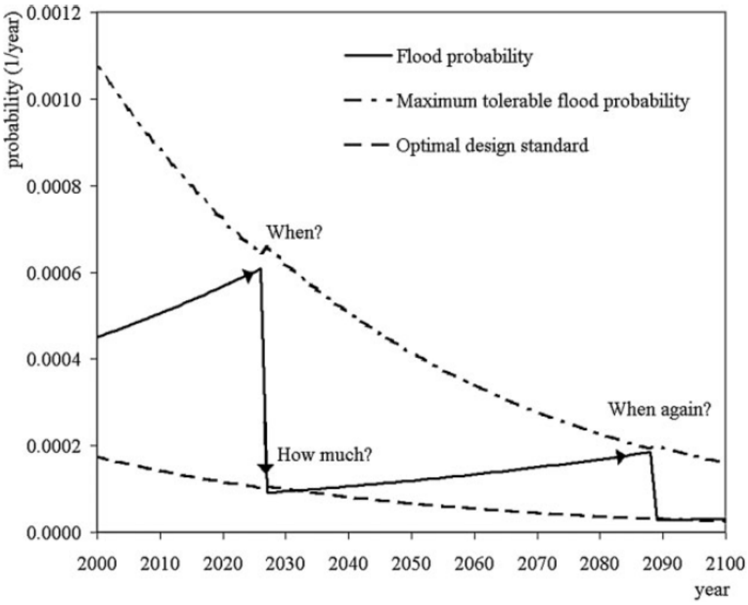
180 This represents the probability that water would exceed the design water level for the dikes surrounding the dike ring areas.

181 JM Kind (2014), "Economically efficient flood protection standards for the Netherlands," *Flood Risk Management*, 7, pp. 103-117.

182 Dutch Water Sector, *Dutch parliament adopts unique risk standards for flood protection*, July 2016.

183 JM Kind (2014), "Economically efficient flood protection standards for the Netherlands," *Flood Risk Management*, 7, pp. 103-117.

Figure 10: Optimal periodical investments



Sources: JM Kind (2014), "Economically efficient flood protection standards for the Netherlands," *Flood Risk Management*, 7, pp. 103 – 117.

The model authors recognized that if households were risk averse (eg, if they were willing to pay a higher "risk premium" than the value of the calculated flood risk reduction), then this approach would result in an underestimate of the true social economic benefits. That is, if risk premiums were factored into the benefits of a flood defense project, then the resulting flood protection standard would be higher than it otherwise would have been.[184]According to the model authors, this does not make a significant difference in the case of the Netherlands, as the government is also the entity to provide compensation for flood damages if they occur. Thus if the compensation is high enough, there is a reduced need to include a risk premium in the cost-benefit analysis.

In a separate study of three communities in Germany, Merz, Elmer and Thieken (2009) incorporated risk aversion into a model that estimated flood damage in these communities and selected an optimal flood risk mitigation strategy for each. The authors used a stylized form of risk aversion, where people were assumed to overweigh events with large damages. Their study found that including this type of risk aversion reduced the benefits of dikes (resistance strategy), and increased the benefits of warning systems (resilience strategy).[185]

184 The authors further found that the risk premium is sensitive to the expected compensation given by the government – increasing from 8% to 41% if the compensation level decreases from 75% to 50%. See JM Kind (2014), "Economically efficient flood protection standards for the Netherlands," *Flood Risk Management*, 7, pp. 103-117.

185 Merz, Elmer and Thieken (2009), "Significance of 'high probability/low damage' versus 'low probability/high damage' flood events," *Natural Hazards and Earth System Sciences*, June 2009.

United Kingdom: Risk-based cost-benefit analysis

The United Kingdom is also prone to floods. In the summer of 2007, severe flooding unexpectedly occurred, impacting around 55,000 properties. An independent review into those floods found that they led to the largest loss of essential services in Great Britain since World War II, leaving almost half a million people without water or electricity. The Pitt Review estimated the insurance industry to pay out over £3 billion in claims, with other costs borne by central government, local public bodies, businesses and private individuals.[186] To manage future risk and uncertainty, the UK government has adopted a different risk-based approach from the Netherlands', one based on cost effectiveness.

The Environment Agency is the entity that oversees flood risk management, assessing the trade-offs between the costs of flood defence mechanisms and the benefits of flooding reduction. Its aim is to minimise the harm of flooding by: (1) reducing the likelihood of flooding; and (2) reducing the impacts when flooding occurs.[187] It has published several long-term planning documents, including the 2014's long-term investment scenarios (LTIS) to manage flood risk and uncertainty in the next 25 years.[188]

The Environment Agency uses a risk-based approach to identify the annual average economically optimal level of flood (and coastal erosion) risk management investment. It assesses "flood damages and the benefits of flood risk management statistically, by multiplying the modelled likelihood of flooding by the flood damage for each type of property".[189] On the risk side, it has found in its 2014 LTIS study that: "There are currently around 2.4 million properties at risk of flooding from rivers and the sea in England. 748,000 of these have at least a 1% annual likelihood of experiencing flooding. About 3 million properties are at risk from surface water flooding in England, around 772,000 of which are at or above the 1% annual likelihood level".[190] On the cost-benefit side, it sought to identify the portfolio of flood defence investments that would maximize the overall net present value (NPV) based on expected costs and benefits. The optimal level of investments is where the NPV can no longer be improved by increasing the investment level (because the benefits associated with one more unit of investment is less than the costs associated with the same increase).[191]

The Environment Agency constructs different scenarios in assessing the annual average optimum investment level.[192] Then, it chooses projects based on a rule of thumb of a benefit cost ratio of five to one on average due to budget constraints (note that this implies a significant

---

186 Michael Pitt (2008), *Floods review*, June 2008.

187 Environment Agency (2009), *Flooding in England: A national assessment of flood risk*, 2009, p. 7.

188 See also the Foresight study in 2004, the Long-term investment strategy in 2009.

189 Environment Agency (2014), *Flood and coastal erosion risk management, Long-term investment scenarios*, 2014, p. 52.

190 Environment Agency (2014), *Flood and coastal erosion risk management, Long-term investment scenarios*, 2014, p. 4.

191 Environment Agency (2014), *Flood and coastal erosion risk management, Long-term investment scenarios*, 2014, p. 16.

192 Environment Agency (2014), *Flood and coastal erosion risk management, Long-term investment scenarios*, 2014, p. 17.

amount of “catch up” investment is required, since many projects have very large benefit cost ratios).[193]

While the risk assessment is said to be risk-neutral, the Environment Agency acknowledges that it expects the actual consequences of flooding to be lower than that modelled because of existing resiliency measures that would reduce the impacts of flooding when it occurs. Examples are: “providing timely and accurate warnings, allowing householders business owners to take action to protect their property, and responding to incidents to reduce the impacts of flooding”.[194] These factors are estimated to reduce economic risk by about 10%.[195]

Takeaways for the NEM

- The overall framework of flood defence planning in the Netherlands aims to reduce flood risk to a specified “acceptable” level, defined as a 1:100,000 risk of flood-related mortality per year.
- Flood defence projects are designed to minimize the costs of meeting the flood protection standard (in expectation). Risk aversion was not incorporated into modelling the flood protection standard because the Government provides compensation to individuals suffering flood damage.
- Flood defence investment in the UK is planned using expected costs and benefits, weighted by probabilities.

193 Environment Agency (2009), *Flooding in England: A national assessment of flood risk*, 2009, p. 11.

194 Environment Agency (2014), *Flood and coastal erosion risk management, Long-term investment scenarios*, 2014, pp. 12-13.

195 Environment Agency (2014), *Flood and coastal erosion risk management, Long-term investment scenarios*, 2014, pp. 12-13.

BOSTON
NEW YORK
SAN FRANCISCO

WASHINGTON
TORONTO
LONDON

MADRID
ROME
SYDNEY

THE Brattle GROUP