Technical Memorandum

AN APPLICATION OF GAME THEORY TO RATE OF RETURN REGULATION IN ALBERTA

by

John P. Evans

Robert E. Evans

VAN HORNE INSTITUTE

University of Calgary

November 2017

EXECUTIVE SUMMARY

Capital-intensive utilities and pipelines are typically subject to rate of return regulation; and regulators are routinely called upon to determine a fair rate of return on assets based on expert evidence. In recent years, the range of expert recommendations proffered by utility and intervenor interests in Alberta have widened to 3-4%, making it increasingly difficult for the Alberta Utilities Commission (Commission) to have confidence in its rate of return decisions or the underlying evidence. The Commission's rate of return decisions have tended to be close to the midpoint of the competing recommendations in generic cost of capital (GCOC) proceedings, thereby reinforcing the wisdom of parties in taking ever more extreme positions in an effort to "pull the average up" or "push the average down." We discuss this frustrating situation in Chapter 1.

Game theory is the field of economics which deals with these kinds of behavioral response situations. To our knowledge, game theory has not played a prominent role in regulatory and public utility economics. Nevertheless, the current regime in which parties are seeking to "pull the average up" or "push the average down" can be interpreted as a "game" as that term is used in economics. And unless the rules in the present "game" are changed, it is difficult to imagine that the parties will embrace a "move to the middle" strategy that reduces the existing gap between the recommendations of utility and intervenor experts.[1]

In Chapter 2, we design specific tests for evaluating alternative games to reward the utilities and the intervenors for moderating their positions respecting rate of return and drawing closer to one another – i.e., "moving to the middle." These tests are applied in Chapter 3, where we created and tested 66 alternative Models using 210 separate tests per Model.[2] We conclude that a superior alternative to the present situation is one in which the parties assume that the Commission will give greatest weight to the recommendation which is closest to the average recommendation and least weight to the recommendation which is furthest from the average, subject to a number of constraints and pre-determined parameters. The specific rules of this Model are set out in Part 3.5.

In Chapter 4, we derive the optimal strategy for each party given the recommended Model and alternative assumptions about the positions that the other party is likely to take. The analysis in Chapter 4 confirms that the optimal, least risky, strategies are associated with moderation rather than extremity. In other words, it pays to "seek out the middle ground."

---

[1] In this paper, the word "game" is used in the technical economic sense and is in no way intended to impugn the motives of the parties or indicate disrespect for the regulatory process or the experts who develop rate of return recommendations. In economics, a "game" is a set of rules that define how parties are expected to behave in a given situation. The "rule" in the current regulatory game as perceived by the parties is that the Commission will "split the difference" between the recommendations and then deduct perhaps 25-50 basis points to arrive at the fair rate of return. In game theory, it does not matter whether such a formal rule exists or not. If the Commission's decisions conform to the rule and if the parties assume that the current *modus operandi* will continue, then they will behave as if the formal rule exists.

[2] We alternatively refer to "Games" as "Models."

In Chapter 5, we ask the tantalizing question *What results would our recommended Model have yielded if it had been applied to the recommendations in each of the seven litigated, multi-company rate of return proceedings by the Commission for major utilities?* The answer is that the average awarded rate of return may not have been materially different from the average of the actual awarded rates of return; however, because of the incentives to "move to the middle," it is likely that the recommendations would have been closer together, giving all Parties a greater confidence in the end result.

The studies, conclusions and recommendations are summarized in Chapter 6 along with suggestions for further study. For sake of brevity, we use a number of defined terms which are capitalized and whose definitions are supplied in the Glossary that follows Chapter 6.

The final conclusion as to what is enough but not too much in the way of return, and rate of return is not precisely supportable on a mathematical basis. If it were, one computer and a few programmers could replace all the regulatory boards in North America, and dispense undeniable justice instantaneously.

National Energy Board · Re Trans-Canada Pipe Lines Limited
Reasons for Decision · December 1971 · Page 6-6

Chapter 1

THE COMMISSION'S FRUSTRATION

Regulators, regulatory practitioners and students of the regulatory process recognise the importance of determining a fair return on the assets committed to capital-intensive utilities and pipelines. The fair return question is typically contentious because: (i) the dollars involved are usually significant; and (ii) as indicated by the National Energy Board in its 1971 *Trans-Canada* decision, there is no mathematical formula which regulatory participants can use to determine the fair return with precision.

The late Dr. Stephen F. Sherwin, a highly-respected regulatory expert, was known to embrace the proposition that opinions respecting the fair return required professional judgment but "not unbridled judgment" – instead, "judgment constrained by the facts." Another regulatory expert cautions, "Such models (i.e., mathematical models) possess a Siren-like quality, lulling the unwary into a blissful and lethargic world of delusory precision."

An obvious consequence of the role of judgment is that independent experts may logically arrive at different recommended rates of return. Recommendation differences may arise due to differences in estimation methods, data, forecasts and risk analyses. In the late 1970s and early 1980s, recommended common equity rates of return from experts whose evidence was sponsored by utilities and those whose evidence was sponsored by intervenors typically differed by approximately 150 basis points (i.e., 1.5%), although differences were sometimes as small as 100 basis points.

Over time, the typical difference between utility and intervenor recommended rates of return increased. From the mid-1980s to 1996, the average difference hovered just below 2.0%. In 1999 and 2000, the typical difference rose to over 2.5%; and in 2001 and 2002, the difference was pushing 3.5%.

In *Decision 2001-96* respecting ATCO Gas (South), the Board stated:

> The Board has reviewed the evidence of Ms. McShane for ATCO and Drs. Booth and Berkowitz for Calgary. The Board is concerned that, despite its volume, the nature of the expert

> evidence provided is ultimately of little probative value to the Board in establishing this important determinant of the utility's revenue requirement.[1]

The Board amplified on the cause of its frustration when it noted in respect of the witnesses' views respecting the market risk premium.

> …these estimates are far enough apart that the underlying evidence is of little value to the Board in establishing an accurate and well justified estimate of the utility rate of return required to maintain the financial integrity of the utility in the eyes of investors and the market. Subsequently, the Board must rely on an examination of past awards to CWNG (Canadian Western Natural Gas) to determine if there is a requirement for adjustments to those awards. The Board is also of the view that alternative methods of determining appropriate utility return may need to be examined for use in future rate cases.[2]

The situation has not improved. Differences in recommended rates of return in the Commission's Generic Cost of Capital (GCOC) proceedings have ranged from 2.4% to 3.8% and have averaged 3.2%. In the most recent GCOC proceeding, the difference between recommendations by witnesses sponsored by the utilities and witnesses sponsored by the intervenors was 3.25%.[3]

In *Decision 20622-D01-2016*, the Commission, focusing on the determination of the appropriate beta factor, expresses a similar frustration to the Board in *Decision 2001-96*. The Commission states:

> In this proceeding, the Commission observes an unusually wide range of recommended betas spanning approximately 470 bps (0.45 to 0.92), which is also substantially larger than the 250 bps span observed in the 2013 GCOC proceeding. The Commission has considered the positions and critiques of all the parties with respect to beta and notes that these positions and critiques are reasonable and generally valid. Consequently, the Commission cannot identify, with any reasonable degree of confidence, a method that allows the Commission to narrow the range of betas recommended by experts in this proceeding.[4]

---

[1] *Decision 2001-96*, page 57.
[2] *Decision 2001-96,* page 58.
[3] See *Decision 20622-D01-2016*. The recommendations that defined the 3.25% range were those of Dr. Villadsen (10.25%) and Dr. Cleary (7.00%).
[4] *Decision 20622-D01-2016*, page 40, paragraph 182.

The Commission's concern about the "wide range of recommended betas" led, in part, to its conclusion that the 2016 decision would focus on *changes* since the 2013 GCOC decision rather than on a *de novo* determination of the appropriate rate of return in 2016. The Commission concludes:

> For the purposes of this decision, the Commission's point of departure is the allowed ROE and approved deemed equity ratios established in the 2013 GCOC decision. From this starting point, the Commission has evaluated the evidence and argument in this proceeding to determine whether changes in the allowed ROE and approved deemed equity ratios from the 2013 GCOC decision are warranted. To that end, the Commission generally considered the directional effect of elements of the evidence and argument in this proceeding on the allowed ROE and approved deemed equity ratios from the 2013 GCOC decision.[5]

To illustrate these trends, we have prepared charts using sample evidence and Board decisions for the 1984 to 2017 period. The test years, sample evidence and Board decisions are described in Table 1.1.

[5] *Decision 20622-D01-2016*, page 6, paragraph 27.

Table 1.1

SAMPLE EVIDENCE AND BOARD DECISIONS

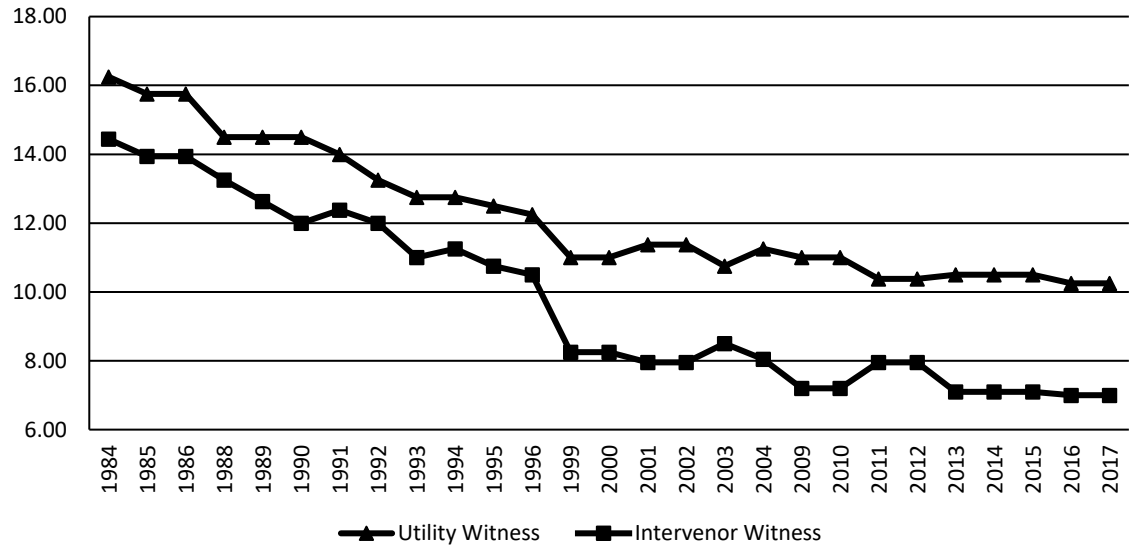
Test Year	Applicant	Company Witness	Intervenor Witness(es)	Board Decision
1984	TransAlta Utilities	Evans	Waters	E84140
1985	TransAlta Utilities	Evans	Waters	E85129
1986	TransAlta Utilities	Evans	Waters	E85129
1988	TransAlta Utilities	Evans	Waters	E89091
1989	TransAlta Utilities	Evans	Waters	E89091
1990	TransAlta Utilities	Evans	Waters	E89091
1991	TransAlta Utilities	Evans	Waters	E91093
1992	Nova Corporation of Alberta	Evans	Waters	E92086
1993	Nova Corporation of Alberta	Evans	Waters	E93060
1994	Nova Gas Transmission	Evans	Waters	E94078
1995	Nova Gas Transmission	Evans	Booth & Berkowitz	U96001
1996	TransAlta Utilities	Evans	Waters & Winter	U97065
1999	Edmonton Power Generation/Transmission	Evans	Waters & Winter	U99099
2000	Edmonton Power Generation/Transmission	Evans	Waters & Winter	U99099
2001	UtiliCorp Networks	Evans	Kryzanowski & Roberts	Settled
2002	UtiliCorp Networks	Evans	Kryzanowski & Roberts	Settled
2003	AltaLink Management	Evans	Booth	2003-061
2004	GCOC 2004	McShane	Kryzanowski & Roberts	2004-052
2009	GCOC 2009	Vilbert	Booth	2009-216
2010	GCOC 2009	Vilbert	Booth	2009-216
2011	GCOC 2011	McShane	Booth	2011-474
2012	GCOC 2011	McShane	Booth	2011-474
2013	GCOC 2015	McShane	Cleary	2191-D01-2015
2014	GCOC 2015	McShane	Cleary	2191-D01-2015
2015	GCOC 2015	McShane	Cleary	2191-D01-2015
2016	GCOC 2016	Villadsen	Cleary	20622-D01-2016
2017	GCOC 2016	Villadsen	Cleary	20622-D01-2016

Sources: Cited evidence and regulatory decisions. We were unable to locate evidence in a major utility litigated proceeding in Alberta for the years 1987, 1997 and 1998. The 2005-2008 awarded rates of return on common equity were the result of formulaic calculations set out in *Decision 2004-052* and did not involve the tendering of expert evidence.

Chart 1.1 shows trends in the utility and intervenor recommended rates of return for each of these witnesses and decisions.

Chart 1.1

Common Equity Rate of Return Recommendations



The widening of the gap between utility and intervenor witness recommendations can be seen visually starting in about 1999. The source of the Commission's recent frustration is apparent from the significant gap which has existed since 2013.

Chart 1.2 shows the same data from a different perspective. The bars indicate the differences between the utility and intervenor expert recommendations for each year.

Chart 1.2

Differences Between Company and Intervenor Witness Common Equity Rate of Return Recommendations

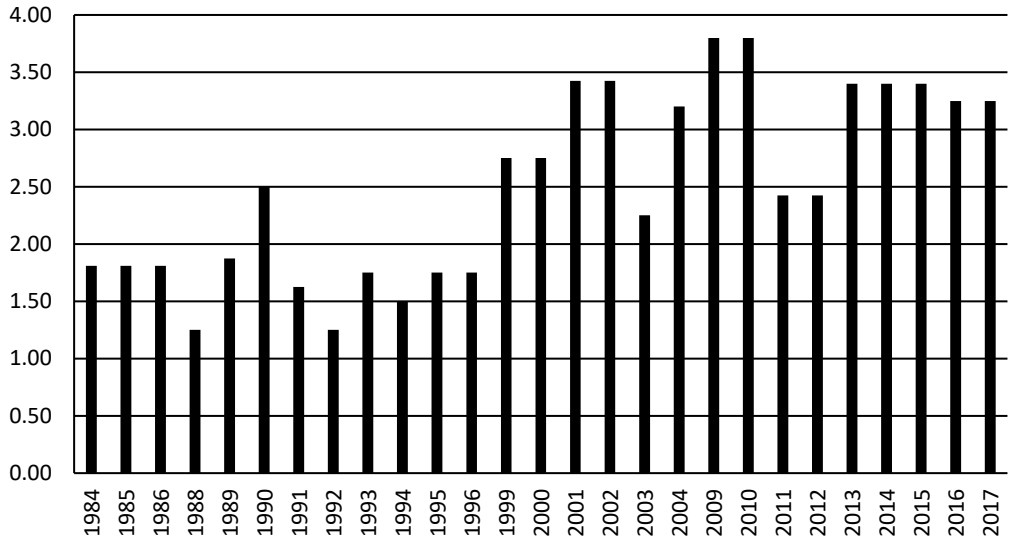
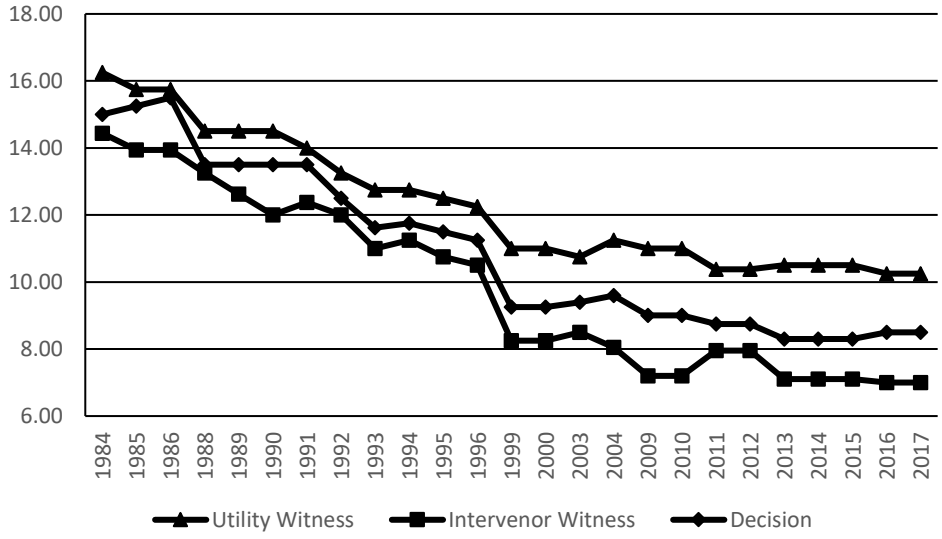


Chart 1.3 shows the same data as Chart 1.1 but with the regulatory rate of return decision added to the graph.

Chart 1.3

Recommended Common Equity Rates of Return and Regulatory Decisions

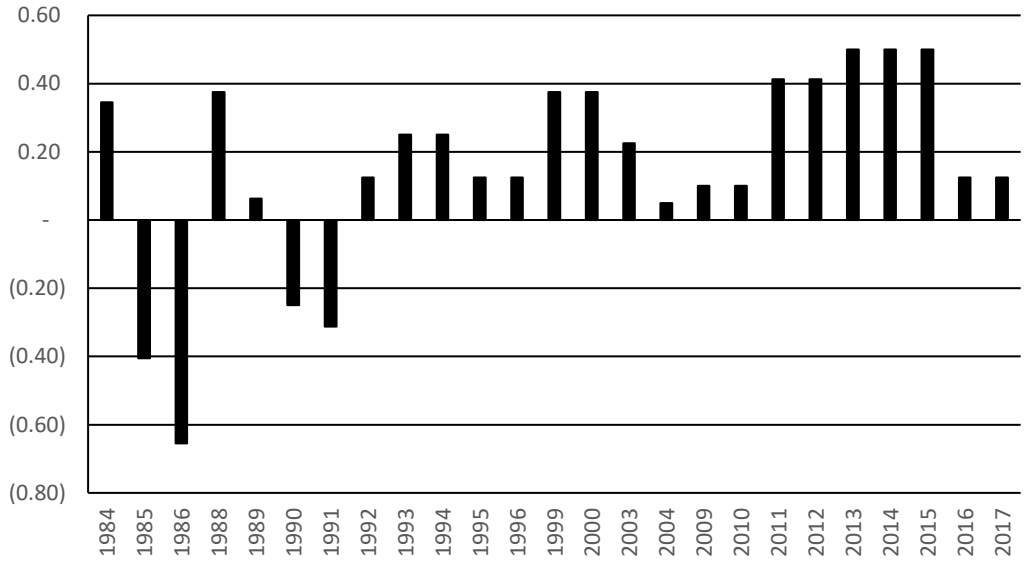


The Commission and its predecessors have made their rate of return findings within the ranges of recommendations proffered by the experts. In the late 1980s, there were years which clearly favoured the recommendations of the utility experts and years which clearly favoured the recommendations of the intervenor experts; but for most years, the Commission's decision was centred between the two sets of recommendations, albeit slightly favouring the intervenor recommendations in recent years.

Chart 1.4 shows the annual differences between the mid-point recommendation of the experts and the regulatory decisions. The scale on the chart is quite small and indicates that, for the most part, Alberta regulators have come close to "splitting the difference."

Chart 1.4

Differences Between Mid-Point Recommendations and Regulatory Decisions



We draw three conclusions from the data in Charts 1.1 – 1.4. First, we agree with the Commission that the wide gap which now exists between utility and intervenor expert recommendations makes it difficult to draw meaningful *de novo* conclusions respecting the appropriate rate of return on common equity.

Second, we note that this recommendation gap was not as wide in the years leading up to, say, 1999.

Third, the tendency of Alberta regulators has been to "split the difference" between the recommendations with the difference between the mid-point of the expert recommendations and the final regulatory decision being 50 basis points or less for the entire study period, with the exception of 1986.

The analysis prompts several questions. Why did the recommendation difference begin to widen in 1999? Why is there still a significant difference in 2017? Is there anything that can be done to shrink the difference in the future?

For most of the years immediately preceding 1999, regulatory decisions were within 25 basis points of the mid-point of the recommendation range. We suggest that a perception may have arisen among utilities and intervenors that whatever their witnesses' recommendations (within reasonable limits), the Commission would simply "split the difference." Therefore, if a utility considered that awards tended to be lean, then it might engage an expert whose recommendation was increasingly higher than what the utility expected to receive in the hope that the expert would "drag the average up." If an intervenor considered that awards tended to be generous, then the intervenor might engage an expert whose recommendation was increasingly lower than what it expected the Commission to grant in the hope that the expert would "drag the average down."

These behaviours by the parties would be perfectly logical; and because of the considerable judgment required in rate of return analysis, experts might well emphasize the upper or lower parts of relevant ranges without violating their intellectual honesty. But as the parties each vied for a better position in "dragging the average up or down," the recommendations diverged and remained divergent.[6]

"Game theory" is the branch of economics which deals with these kinds of behavioral situations. Perhaps the most famous problem in game theory is The Prisoner's Dilemma. Two men are accused of robbing a bank. The police capture both and put them in separate rooms. Each suspect is told that if he confesses and implicates the other suspect, then he will go to jail for five years. If he refuses to implicate the other suspect, then he will either go to jail for ten years if he is implicated by the other suspect; or he will get off with no jail time if both suspects refuse to implicate each other. Game theory demonstrates that the optimal solution is for both prisoners to confess.

Game theory economists are routinely employed by Canada's Department of National Defence, the U. S. Department of Defense and others to explore strategic questions of national defense. For example, if the U. S. launched a first-strike nuclear attack against North Korea, what would be the likely response from North Korea or China or others? A game theory economist would study this "game" and provide an answer given the available intelligence about the other "players."

To our knowledge, game theory has yet to play a prominent role in regulatory and public utility economics. But the current situation begs for consideration and a possible solution from the world of game theory, because the utilities, the intervenors and the Commission are all involved in a "game" where the parties before the Commission are attempting to

[6] We reject as disingenuous and contrary to reality the notion that utilities and intervenors are unconcerned with the truth and employ experts who are intellectually dishonest. We regard this notion of the world as naïve. Nevertheless, we do recognize the realistic possibility that parties and their experts may not be encouraged strongly enough to move to more central positions and that parties will naturally tend to retain independent experts whose analyses lead to recommendations that are consistent with the thinking of the retaining party.

influence the Commission's decision by taking increasingly more divergent positions; and the Commission is, for the most part, perceived to be maintaining a "split the difference" approach which, in turn, encourages the parties to continue playing the current "game" *in perpetuum*.[7]

Unless the rules in the present "game" are changed, the parties are unlikely to embrace a "move to the middle" strategy that will reduce the existing gap between the expert recommendations of utility and intervenor witnesses. What then, if anything, can be done to change the "rules of the game" so that incentives and motivations encourage parties to moderate their positions, thereby improving the usefulness of expert evidence for regulatory purposes? Is there a solution in game theory? Even if there is not an optimal set of "rules," can the "rules of the game" be changed so that there is no longer an incentive to take increasingly-extreme positions?

We set out to research whether there is a game theory solution that will improve Alberta's GCOC regulatory process. We doubt that there is a single-best optimal solution; however, we have identified a game from among many that we tried which offers superior incentives and, hopefully, outcomes.

We offer a final caveat about the requirements for a fair return. The Supreme Court of Canada in *Northwestern Utilities Ltd. v. City of Edmonton*[8] established the well-accepted triumvirate of fair return criteria that are applied by utility regulators. The fair return must be sufficient to:

- Maintain the financial integrity of the utility (the "financial integrity standard")
- Permit the utility to attract capital on reasonable terms (the "capital attraction standard")
- Be consistent with returns being offered on other investments exposed to a similar level of risk (the "comparable earnings standard")

There is no assurance that any game theory solution will competently address itself to these criteria, because that solution will be subject to the same vagaries and estimation problems as the current regime. And yet, regulators must respect the legal requirements for a fair return. The solution is to reset the Parties' expectations regarding how the Commission will determine what we refer to as the Starting Point Rate of Return

[7] In this paper, the word "game" is used in the technical economic sense and is in no way intended to impugn the motives of the parties or indicate disrespect for the regulatory process or the experts who develop rate of return recommendations. In economics, a "game" is a set of rules that define how parties are expected to behave in a given situation. The "rule" in the current regulatory game as perceived by the parties is that the Commission will "split the difference" between the recommendations and then deduct perhaps 25-50 basis points to arrive at the fair rate of return. In game theory, it does not matter whether such a formal rule exists or not. If the Commission's decisions conform to the rule and if the parties assume that the current *modus operandi* will continue, then they will behave in response as if the formal rule exists.

[8] 2 DLR 4 [1929]

(SPROR). The SPROR is the rate of return broadly indicated by the recommendations on the record – i.e., the "starting point" for the Commission's consideration of the fair return question. And if the Commission finds that the SPROR satisfies the fair return criteria and should therefore be accepted as the fair rate of return, then parties will logically alter their behaviors to maximize or minimize the SPROR.

The answer to the Commission's present frustration lies in defining a game in which the SPROR-maximizing strategy adopted by utilities is to table recommended rates of return at the lower end of their reasonable range and in which the SPROR-minimizing strategy adopted by intervenors is to table recommended rates of return at the higher end of their reasonable range. The quest to define or create the rules of this game is the subject of Chapter 2.

Chapter 2

TESTING AND EVALUATING ALTERNATIVE GAMES

In Chapter 2, we consider 66 alternative models or games designed to reward the utilities and the intervenors for moderating their positions respecting rate of return – i.e., "moving to the middle." The current rule perceived by the parties for determining the Starting Point Rate of Return (SPROR) is something like "average the most extreme recommendations and then deduct 25-50 basis points to arrive at the fair rate of return." This rule has logically led recommendations to diverge significantly and is the reason for embarking on this journey to find new rule(s) that will produce a different result.

Chapter 2 is brimming with shorthand, defined terms, which are described in the Glossary at the end of this Technical Memorandum. When a term is first used, we have included the definition in a footnote or in the accompanying text. Chapter 2 is divided into the following parts.

2.1 Utilities, Intervenors, Parties, Games, Models, Strategies, Outcomes, Strategy Combinations, Reality Gaps, Strategy Gaps, Environments, Recommendation Spreads and Internal Differences

In Part 2.1, we provide key definitions that are required to begin the discussion of the testing and analyses which follow.

"Utilities" refers to companies whose earnings are subject to rate of return regulation. Under most circumstances, Utilities have an incentive to maximize the rate of return awarded by the regulator.

"Intervenors" refers to entities whose interests are generally furthered by minimizing the rate of return awarded by the regulator. Customer groups and consumer advocates are examples of prominent Intervenors.

"Parties" refers to Utilities and Intervenors collectively.

A "Game" in the present context is a set of rules that define how the Starting Point Rate of Return (SPROR) will be determined.[1] The words "Model" and "Game" are synonymous.

"Strategies" are names given to recommended rates of return that are set at specified distances from the rate of return that Utilities and Intervenors privately forecast that the Commission is most likely to award. The Strategy names are Forecast, Strategic, Moderate, Immoderate and Extreme. For example, suppose that the Utilities privately forecast that the Commission will award an 8.5% common equity rate of return in an upcoming Generic Cost of Capital (GCOC) proceeding, then the Forecast rate of return is 8.5%. If the Strategy Gap (see definition below) is 0.5%, then the Strategic rate of return is 9.0%. The Moderate rate of return is 9.5%. The Immoderate rate of return is 10.0%; and the Extreme rate of return is 10.5%. Suppose that the Intervenors also privately forecast that the Commission is likely to award an 8.5% rate of return. Assuming the same Strategy Gap (i.e., 0.5%), the possible Intervenor Strategies are Forecast (8.5%), Strategic (8.0%), Moderate (7.5%), Immoderate (7.0%) and Extreme (6.5%).

"Outcomes" refers to a set of two Utility Strategies and two Intervenor Strategies based on the recommendations of two Utility experts and two Intervenor experts.[2] For example, a possible Outcome is Extreme/Immoderate (Utility) and Immoderate/Strategic (Intervenor).[3]

"Strategy Combinations" refers to two Strategies adopted by the same Party. To illustrate, Extreme/Extreme is a Strategy Combination that could be adopted by either the Utilities or the Intervenors. Moderate/Strategic is an alternative Strategy Combination.

---

[1] The Starting Point Rate of Return (SPROR) is the rate of return broadly indicated by the recommendations on the record – i.e., the "starting point" for the Commission's consideration of the fair return question.

[2] In the five GCOC hearings commencing in 2004, the Intervenors have consistently sponsored two rate of return experts. In contrast, the Utilities sponsored four experts in 2004, three experts in 2009, one expert in each of 2011 and 2015 and two experts in 2016. Our work assumes that both Parties will sponsor two expert recommendations. If a Party chooses to sponsor more than two expert recommendations, then the most extreme recommendations (i.e., the highest recommendation for the Utilities and the lowest recommendation for the Intervenors) will be excluded to reduce the number of recommendations to two. If a Party chooses to proffer only one recommendation, then that recommendation will be "counted twice" so that each Party will have associated with it two recommendations. The fact that a recommendation is excluded as being "extreme" does not mean that the Commission will ignore the evidence. It simply means that the recommendation will not be considered in developing the Starting Point Rate of Return (SPROR) described in Chapter 1. As a result, the Parties may conclude that they should avoid making too many recommendations and that they should avoid extreme recommendations.

[3] Henceforth, the convention for describing Outcomes will place the two Strategies for the Utilities first followed by the two Strategies for the Intervenors. Thus, the Outcome described in this sentence would be referred to as Extreme/Immoderate/Immoderate/Strategic.

"Reality Gap" refers to any difference that may exist between the private perceptions of Utilities and Intervenors respecting the common equity rate of return which the Commission will ultimately award (i.e., the rate of return associated with each Party's Forecast strategy). To illustrate, if the Utilities believe that the Commission is likely to award an 8.75% rate of return and if the Intervenors believe that the Commission is likely to award an 8.25% rate of return, then the Reality Gap is 0.5% (= 8.75% *less* 8.25%). The Parties are likely to develop their perceptions of reality from the last rate of return awarded by the Commission and information on changes in interest rates, bond yields, inflation and other general indicators of trends in the level of capital cost. If the Parties have the same private perception of what the Commission's decision will ultimately be, then the Reality Gap is 0%. In our work, we consider Reality Gaps of 0% and 1.0%. Reality Gaps of greater than 1.0% would imply a naivete that we consider unlikely for either Utilities or Intervenors.

"Recommendation Spread" refers to the difference between the *highest* Utility recommendation and the *lowest* Intervenor recommendation – i.e., the range which encompasses *all* recommendations. We also refer to the "Internal Recommendation Spread," which is defined as the difference between the *lowest* Utility recommendation and the *highest* Intervenor Recommendation.

"Strategy Gap" is referred to in the definition of "Strategies." The Strategy Gap is the difference between common equity rates of return as Parties move from one Strategy to the next. In this paper, we consider Strategy Gaps of 0.25% and 0.5%.

"Environments" refers to three specific combinations of Reality Gaps and Strategy Gaps – a Reality Gap of 0% with a Strategy Gap of 0.25%, a Reality Gap of 1% with a Strategy Gap of 0.25% and a Reality Gap of 0% with a Strategy Gap of 0.5%. The significance of these choices is that they represent Maximum Recommendation Spreads of 2.0%, 3.0% and 4.0% respectively.[4] As shown on Chart 1.2, Recommendation Spreads have been in the range of approximately 2.0-4.0% for at least the past 20 years and within the range of 1.5-4.0% for virtually the entire 1984-2017 study period. Each Environment has associated with it a set of ten rates of return (= 2 Parties x 5 Strategies). For example, Table 2.1 shows the ten rates of return associated with the Environment defined by a Reality Gap of 1.0% and a Strategy Gap of 0.25%.[5]

[4] The Recommendation Spread is the difference between rates of return recommended by Utilities and rates of return recommended by Intervenors.

[5] For analytical purposes, we assume throughout that if the Parties are in an Environment with a 0% Reality Gap, then the Forecast rate of return for both Utilities and Intervenors will be 8.50%, the 2017 GCOC rate of return. The selection of this value is arbitrary; and the qualitative conclusions respecting the efficacy of different Models are independent of this value.

Table 2.1

STRATEGIES AND RATES OF RETURN ASSOCIATED WITH A REALITY GAP OF 1.0% AND A STRATEGY GAP OF 0.25%

Utilities	
Extreme	10.00%
Immoderate	9.75
Moderate	9.50
Strategic	9.25
Forecast	9.00
Reality Gap of 1.0%	
Intervenors	
Forecast	8.00%
Strategic	7.75
Moderate	7.50
Immoderate	7.25
Extreme	7.00

Thus, the "Extreme" position for the Utilities is the *highest* rate of return moving away from what the Utilities forecast the end result is likely to be. In contrast, the "Extreme" position for the Intervenors is the *lowest* rate of return moving away from the Intervenors' forecast end result. The "Recommendation Spread" varies from 1.0% (Forecast/Forecast) to 3.0% (Extreme/Extreme) in this Environment.

"Internal Difference" refers to the percent difference between the recommendations of the two Utility-sponsored experts and the percent difference between the recommendations of the two Intervenor-sponsored experts. For example, if the two Utility experts recommend rates of return of 9.75% and 9.25%, then the Utility Internal Difference is 0.5% (= 9.75% *less* 9.25%). If the two Intervenor experts recommend rates of return of, say, 7.75% and 7.5%, then the Intervenor Internal Difference is 0.25% (= 7.75% *less* 7.5%).

"Internal Difference" can also be expressed as the number of Strategies separating the two recommendations of either the Utilities or the Intervenors. For example, if the recommendations of the Utilities are Extreme and Moderate, then the Internal Difference is 2 (i.e., Extreme to Immoderate and Immoderate to Moderate). If the recommendations of the Intervenors are Moderate and Strategic, then the Internal Difference is 1. An Internal Difference of, say, 0/2 means that the Utilities have an Internal Difference of 0; and the Intervenors have an Internal Difference of 2 (e.g., Immoderate/Immoderate/ Immoderate/Strategic). An Internal Difference of, say, 2/1 means that the Utilities have an Internal Difference of 2; and the Intervenors have an Internal Difference of 1 (e.g., Extreme/Moderate/Moderate/ Strategic).

2.2 Vickrey Auctions, Incentives to Reveal the Truth and Closed Bids

In 1996, Canadian economist William Vickrey won the Nobel Memorial Prize in Economics for his work in applying game theory to auctions. Vickrey is best-remembered for his work on what is known as the Vickrey Auction – a sealed-bid auction in which the highest bidder wins the auction but pays the second-highest price. The optimal strategy for bidders in a Vickrey Auction is to bid a price that reflects the bidder's perception of true value. By bidding a price that reflects "true value," each participant maximizes his opportunity to win the auction but is assured that he will not necessarily be paying a "premium" above what is required to win.

In *Algorithms to Live By: The Computer Science of Human Decisions*, authors Brian Christian and Tom Griffiths explain the Vickrey Auction this way.

> We've seen how seemingly innocuous auction mechanisms, for instance, can run into all sorts of problems: overthinking, overpaying, runaway cascades. But the situation is not completely hopeless. In fact, there's one auction design in particular that cuts through the burden of mental recursion like a hot knife through butter. It's called the Vickrey auction. Named for Nobel Prize-winning economist William Vickrey, the Vickrey auction, just like the first-price auction is a 'sealed bid' auction process. That is, every participant simply writes down a single number in secret, and the highest bidder wins. However, in a Vickrey auction, the winner ends up paying not the amount of their own bid, but that of the *second* place bidder. That is to say, if you bid $25 and I bid $10, you win the item at *my* price: you only have to pay $10. To a game theorist, a Vickrey auction has a number of attractive properties. And to an algorithmic game theorist in particular, one property especially stands out: the participants are incentivized to be honest. In fact, there is no better strategy than just bidding your 'true value' for the item – exactly what you think the item is worth. Bidding any more than your true value is obviously silly, as you might end up stuck buying something for more than you think it's worth. And bidding any less than your true value (i.e., shading your bid) risks losing the auction for no good reason, since it doesn't save you any money – because if you win, you'll only be paying the value of the second-highest bid, regardless of how high your own was. This makes the Vickrey auction what mechanism designers call 'strategy-proof,' or just 'truthful.' In the Vickrey auction, honesty is literally the best policy. Even better, honesty remains the best policy regardless of whether the other bidders are honest themselves. In the prisoner's dilemma, we saw how defection turned out to be the 'dominant' strategy – the best move no matter whether your partner defected or cooperated. In a Vickrey auction, on the other hand, honesty is the dominant strategy. This is the mechanism designer's holy grail. You do not need to strategize or recurse. Now, it seems like the Vickrey auction would cost the seller some money compared to the first-price auction, but this isn't necessarily true. In a first-price auction, every bidder is shading

> their bid down to avoid overpaying; in the second-price Vickrey auction, there's no need to – in a sense, the auction itself is optimally shading their bid *for them*. In fact, a game-theoretic principle called 'revenue equivalence' establishes that over time, the average expected sale price in a first-price auction will converge to precisely the same as in a Vickrey auction. Thus, the Vickrey equilibrium involves the same bidder winning the item for the same price – without any strategizing by any of the bidders whatsoever. As Tim Roughgarden tells his Stanford students, the Vickrey auction is 'awesome.'[6]

The incentive to "tell the truth" makes the Vickrey Auction appealing. Brinkmanship is not rewarded. Truth is the optimal strategy. Potential buyers do not have to worry about what others might do, they know that their best strategy is to bid their highest price.

The situation facing the Commission and the Parties is not dissimilar. The Commission is seeking truth; but the current Model by which the Parties perceive that the Commission makes its decisions is not one which encourages the Parties to reveal their true views on the appropriate rate of return. Philosophically, our work seeks to provide the Commission with a type of Vickrey Auction that drives Parties to proffer more central or moderate recommendations that are more reflective of "truth," thereby reducing Recommendation Spreads and providing the Commission with greater confidence respecting the appropriate SPROR.

To our knowledge, no academic work has been done on designing a Vickrey Auction whose purpose is to "drive parties to the centre" rather than "drive parties to reveal their highest bids." We make no pretense of having created a generic solution that mathematically solves the former. Our work does, however, create solutions that are superior to the current Alberta regulatory Model in "driving parties to the centre."

A final comment is that for any solution to even approximate the efficacy of the Vickrey Auction, the Parties must participate in a "sealed-bid auction." In other words, neither the Utilities nor the Intervenors can have an informational advantage in the sense of knowing what the other Party's experts are recommending. Thus, any Vickrey-like Model that improves the Alberta regulatory process and encourages truth-telling by the Parties must be based on simultaneous filings in GCOC proceedings. Otherwise, the Intervenors will have an informational advantage if the Utilities file first; or the Utilities will have an informational advantage if the Intervenors file first.

Inasmuch as the GCOC is a generic proceeding and does not, of itself, change the rates and tariffs of the Utilities, we consider that it would be wholly appropriate for the Commission to consider that there are no "applicants" – only Parties (both Utilities and Intervenors) who would be required to simultaneously file their evidence (or at least reveal their recommendations at the same time) so that a Vickrey-style auction process can be implemented.

---

[6] Christian, Brian and Tom Griffiths, *Algorithms to Live By: The Computer Science of Human Decisions*, (Penguin Canada: Toronto), 2016, pages 252-253. Brian Christian is a well-known writer. Tom Griffiths is a Professor of Cognitive Science at the University of California-Berkeley.

2.3 The Test Grid

The alternative Models in this study are tested to ensure that they generally reward moderation and penalize extremity, thus "driving Parties to the middle." The testing process for each Model requires an examination of the possible Outcomes in each of the three environments. The Test Grid described here identifies the specific Outcomes that are tested.

For a particular Environment, each of the four recommendations (two for the Utilities and two for the Intervenors) is associated with one of five Strategies (Forecast, Strategic, Moderate, Immoderate or Extreme) so that each of the Utilities and the Intervenors has 15 possible Strategy Combinations.[7] Thus, there are 225 (= 15 x 15) Outcomes given that each of the Utilities and the Intervenors can proffer recommendations which are consistent with 15 Strategy Combinations.

It would be possible to test each Model against all 225 Outcomes under each of the three Environments. But inasmuch as we will be testing 66 Models (discussed in Part 2.4 below), the resulting 44,550 tests (= 225 Outcomes x 66 Models x 3 Environments) are probably unnecessary and would be unduly burdensome. To reasonably limit the testing, the following Outcomes have been excluded.

> **Exclusion 1.** The 125 Outcomes in which one or both of the Utilities or the Intervenors proffers a Forecast recommendation have been excluded. It is unlikely that either Party would sponsor the evidence of an expert whose recommendation is exactly what that Party believes will eventually be granted by the Commission. The probability of these Outcomes is sufficiently remote that they are set aside for testing purposes.[8]
>
> **Exclusion 2.** An additional 30 Outcomes are excluded where the Utility and Intervenor recommendations have the same Internal Difference, which mathematically leads to a simple average of the four recommendations under each of the 66 Models and which therefore will not affect any test of the relative efficacies of the Models.[9]

---

[7] Five strategies taken two at a time *plus* five combinations where the same strategy is adopted in respect of both recommendations (e.g., Extreme/Extreme or Strategic/Strategic). Mathematically: (5! / 2! 3!) + 5 = 15.

[8] The 125 Outcomes may be calculated by deducting from 225 the number of Outcomes that use only Extreme, Immoderate, Moderate and Strategic. Mathematically: (4! / 2! 2!) + 4 = 10 for each of the Utilities and the Intervenors. Thus, there are 100 (= 10 x 10) Outcomes that do not include a Forecast recommendation. Stated otherwise, the number of Outcomes excluded by virtue of assuming that a Party will table a recommendation consistent with their Forecast is 125 (= 225 *less* 100).

[9] The 30 additional Outcomes are the four 0/0 Internal Difference Strategy Combinations for the Utilities (i.e., Extreme/Extreme, Immoderate/Immoderate, Moderate/Moderate and Strategic/Strategic) multiplied by the four 0/0 Strategy Combinations for the Intervenors *plus* the three 1/1 Strategy Combinations for the Utilities (i.e., Extreme/Immoderate, Immoderate/Moderate and Moderate/Strategic) multiplied by the three 1/1 Strategy Combinations for the Intervenors *plus* the two 2/2 Strategy Combinations for the Utilities (i.e., Extreme/Moderate and Immoderate/Strategic) multiplied by the two 2/2 Strategy Combinations for the Intervenors *plus* the one 3/3 Strategy Combination for the Utilities (i.e., Extreme/Strategic) multiplied by the one 3/3 Strategy Combination for the Intervenors.

There are 70 remaining Outcomes (= 225 *less* 125 excluded under Exception 1 *less* 30 excluded under Exception 2). Thus, we undertook 13,860 tests (= 70 Outcomes x 66 Models x 3 Environments).

Table 2.2 is a test grid which shows each of the 225 possible Outcomes, those which are excluded from testing under Exclusions 1 and 2 and the remaining 70 Outcomes on which we focus.

Table 2.2

TEST GRID VISUALLY INDICATING OUTCOMES USED TO EVALUATE ALTERNATIVE MODELS FOR DETERMINING STARTING POINT RATE OF RETURN

		Intervenor Strategies														
Utility Strategies		Extreme Extreme	Extreme Immoderate	Extreme Moderate	Extreme Strategic	Extreme Forecast	Immoderate Immoderate	Immoderate Moderate	Immoderate Strategic	Immoderate Forecast	Moderate Moderate	Moderate Strategic	Moderate Forecast	Strategic Strategic	Strategic Forecast	Forecast Forecast
Extreme	Extreme	Exclusion 2	Test 1	Test 2	Test 3	Exclusion 1	Exclusion 2	Test 4	Test 5	Exclusion 1	Exclusion 2	Test 6	Exclusion 1	Exclusion 2	Exclusion 1	Exclusion 1
Extreme	Immoderate	Test 36	Exclusion 2	Test 7	Test 8	Exclusion 1	Test 9	Exclusion 2	Test 10	Exclusion 1	Test 11	Exclusion 2	Exclusion 1	Test 12	Exclusion 1	Exclusion 1
Extreme	Moderate	Test 37	Test 42	Exclusion 2	Test 13	Exclusion 1	Test 14	Test 15	Exclusion 2	Exclusion 1	Test 16	Test 17	Exclusion 1	Test 18	Exclusion 1	Exclusion 1
Extreme	Strategic	Test 38	Test 43	Test 48	Exclusion 2	Exclusion 1	Test 19	Test 20	Test 21	Exclusion 1	Test 22	Test 23	Exclusion 1	Test 24	Exclusion 1	Exclusion 1
Extreme	Forecast	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1
Immoderate	Immoderate	Exclusion 2	Test 44	Test 49	Test 54	Exclusion 1	Exclusion 2	Test 25	Test 26	Exclusion 1	Exclusion 2	Test 27	Exclusion 1	Exclusion 2	Exclusion 1	Exclusion 1
Immoderate	Moderate	Test 39	Exclusion 2	Test 50	Test 55	Exclusion 1	Test 60	Exclusion 2	Test 28	Exclusion 1	Test 29	Exclusion 2	Exclusion 1	Test 30	Exclusion 1	Exclusion 1
Immoderate	Strategic	Test 40	Test 45	Exclusion 2	Test 56	Exclusion 1	Test 61	Test 63	Exclusion 2	Exclusion 1	Test 31	Test 32	Exclusion 1	Test 33	Exclusion 1	Exclusion 1
Immoderate	Forecast	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1
Moderate	Moderate	Exclusion 2	Test 46	Test 51	Test 57	Exclusion 1	Exclusion 2	Test 64	Test 66	Exclusion 1	Exclusion 2	Test 34	Exclusion 1	Exclusion 2	Exclusion 1	Exclusion 1
Moderate	Strategic	Test 41	Exclusion 2	Test 52	Test 58	Exclusion 1	Test 62	Exclusion 2	Test 67	Exclusion 1	Test 69	Exclusion 2	Exclusion 1	Test 35	Exclusion 1	Exclusion 1
Moderate	Forecast	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1
Strategic	Strategic	Exclusion 2	Test 47	Test 53	Test 59	Exclusion 1	Exclusion 2	Test 65	Test 68	Exclusion 1	Exclusion 2	Test 70	Exclusion 1	Exclusion 2	Exclusion 1	Exclusion 1
Strategic	Forecast	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1
Forecast	Forecast	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1	Exclusion 1

Exclusion 1 · The 125 Outcomes in which one or both of the Utilities or the Intervenors proffers a Forecast recommendation.

Exclusion 2 · The 30 additional Outcomes where the Utility and Intervenor recommendations have the same Internal Difference.

2.4 Defining 66 Models for Testing

We developed 66 Models which we consider to be worthy of testing. Twenty-four Models are referred to as Initial Models, because they capture our first thoughts on alternative ways for determining the SPROR. Based on our work with the Initial Models, we identified a logic problem that dramatically reduced the efficacy of the Initial Models. We addressed the logic problem in the next 24 Models, which we refer to as Weighted Average Fulcrum Adjustment (WAFA) Models.

Finally, we recognised that there is a degree of arbitrariness in determining some of the weights attached to the various recommendations in the WAFA Models. To address this potential concern, we tested a series of 18 Models that have Internally-Calculated Exponential (ICE) weights and which are referred to as ICE Models. The ICE Models have many of the same features as WAFA Models and are yet another extension of our thinking about how to improve the incentives for Parties to "move to the middle."

In this Part 2.4, we describe the 24 Initial Models, the logic problem, the 24 WAFA Models, the arbitrary weighting concern and the 18 ICE Models. But first, it is necessary to define Structures, the Simple Average Rule and Internal Difference Adjustments.

Structures, the Simple Average Rule and Internal Difference Adjustments

"Structure" refers to the basic method of calculation used in each of the Initial Models. We developed six alternative Structures which are described later in this Part 2.4. To illustrate, the simplest Structure is to take that recommendation closest to the average of the four recommendations and adopt that closest recommendation as the SPROR. An example of a more complex Structure is to calculate the SPROR as the weighted average of the four recommendations with the recommendation closest to the average receiving a weight of 4x, the recommendation furthest from the average receiving a weight of 1x and the other two recommendations receiving a weight of 2x.

Half of the Initial Models are subject to what we refer to as the Simple Average Rule. The Simple Average Rule states that if the Internal Recommendation Spread is less than or equal to a pre-specified percentage, then the SPROR is the simple average of those two values. In other words, if the differences between the Parties are fairly modest, then there is no need for more complex methods of calculation; and the current regime of "splitting the difference" will be reasonably satisfactory. The alternative pre-specified percentages that we consider in our research are 1.0% and 0.0% (i.e., no Simple Average Rule); and these values are referred to as Simple Average Limits.

Each of the Utilities and the Intervenors have some control over their respective Internal Differences. They will certainly be aware of the Internal Differences of experts in their group even before the evidence is filed. The Commission should rightly expect that Parties will do their best to avoid large Internal Differences which will only make the Commission's task more difficult and the evidence less credible.

To encourage the Parties to avoid large Internal Differences, some of the Initial Models make an explicit adjustment which reflects the Internal Differences of the Utilities and the Intervenors, conferring a relative benefit on the Party with the smallest Internal Difference. These adjustments are referred to as Internal Difference Adjustments.

The following are descriptions of the 24 Initial Models.

Initial Models

Model 1 · The SPROR is the recommendation closest to the average of the four recommendations · No Internal Difference Adjustment · Simple Average Limit = 1%

Model 2 · The SPROR is the average recommendation excluding the recommendation farthest from the average of the four recommendations · No Internal Difference Adjustment · Simple Average Limit = 1%

Model 3 · The Party whose recommendation is closest to the average of the four recommendations has both of its recommendations reflected in an average that excludes the recommendation of the other Party that is furthest from the average · No Internal Difference Adjustment · Simple Average Limit = 1%

Model 4 · The SPROR is a weighted average where the recommendation closest to the average is given 4x weight, the recommendation farthest from the average is given 1x weight and the other two recommendations are given 2x weight · No Internal Difference Adjustment · Simple Average Limit = 1%

Model 5 · The SPROR is a weighted average where the recommendation closest to the average is given 2x weight, the recommendation farthest from the average is given no weight and the other two recommendations are given 1x weight · No Internal Difference Adjustment · Simple Average Limit = 1%

Model 6 · The SPROR is the average recommendation excluding those recommendations which are more than one standard deviation from the average · No Internal Difference Adjustment · Simple Average Limit = 1%

Model 7 · The SPROR is the recommendation closest to the average of the four recommendations · Includes Internal Difference Adjustment · Simple Average Limit = 1%

Model 8 · The SPROR is the average recommendation excluding the recommendation farthest from the average of the four recommendations · Includes Internal Difference Adjustment · Simple Average Limit = 1%

Model 9 · The Party whose recommendation is closest to the average of the four recommendations has both of its recommendations reflected in an average that excludes the recommendation of the other Party that is furthest from the average · Includes Internal Difference Adjustment · Simple Average Limit = 1%

Model 10 · The SPROR is a weighted average where the recommendation closest to the average is given 4x weight, the recommendation farthest from the average is given 1x weight and the other two recommendations are given 2x weight · Includes Internal Difference Adjustment · Simple Average Limit = 1%

Model 11 · The SPROR is a weighted average where the recommendation closest to the average is given 2x weight, the recommendation farthest from the average is given no weight and the other two recommendations are given 1x weight · Includes Internal Difference Adjustment · Simple Average Limit = 1%

Model 12 · The SPROR is the average recommendation excluding those recommendations which are more than one standard deviation from the average · Includes Internal Difference Adjustment · Simple Average Limit = 1%

Models 13 – 24 are the same as Models 1 – 12 respectively, except that the Simple Average Limit for Models 13-24 is 0% rather than 1%.

Consider the following three examples that illustrate several of the Initial Models.

Example 1. Assume that the Utility witnesses recommend rates of return of 9.0% and 9.5%; and the Intervenor witnesses recommend rates of return of 7.0% and 7.25%. Applying, say, Model 1 to this situation, the average recommendation is 8.19%; and the differences between each of the recommendations and the average are respectively 0.81%, 1.31%, 1.19% and 0.94%. Since Model 1 determines the SPROR as the recommendation closest to the average, the SPROR is 9.0%, because 0.81% is the smallest difference from the average. The Simple Average Rule is not operable, because the 1.75% difference between the lowest Utility recommendation and the highest Intervenor recommendation (= 9.0% *less* 7.25%) is greater than the 1.0% Simple Average Limit for Model 1.

Example 2. Assume that the Utility witnesses recommend rates of return of 9.0% and 9.25%; and the Intervenor witnesses recommend rates of return of 7.5% and 8.0%. Applying, say, Model 6, the difference between the lowest Utility and the highest Intervenor recommendations is 1.0% (= 9.0% *less* 8.0%). The Simple Average Limit for Model 6 is 1.0%. Thus, the Simple Average Rule applies; and the SPROR is 8.5% - the simple average of 9.0% and 8.0%.

Example 3. Assume that Model 23 is applied to Utility witness recommendations of 9.5% and 10.0% and Intervenor witness recommendations of 7.75% and 8.0%. The Simple Average Rule does not apply, because the Simple Average Limit has been set to 0%. Model 23 determines the SPROR as a weighted average with 2x weight given to the recommendation closest to the average, zero weight given to the recommendation farthest

from the average and 1x weight given to the other two averages. An adjustment for Internal Differences is made in the calculations. The average of the four recommendations is 8.81%. The raw differences between each of the four recommendations and the average are 0.69%, 1.19%, 1.06% and 0.81% respectively. To make the Internal Difference Adjustment, we add the Internal Differences of 0.5% for the Utilities and 0.25% for the Intervenors to these raw differences. The adjusted differences are therefore 1.19%, 1.69%, 1.31% and 1.06%. The smallest adjusted difference (1.06%) is associated with the 8.0% Intervenor recommendation; and so that recommendation receives 2x weight in the calculation of the SPROR. The largest adjusted difference (1.69%) is associated with the 10.0% Utility recommendation; and so that recommendation receives zero weight. The 9.5% and the 7.75% each receive a weight of 1; and the weighted average SPROR is 8.31% (=((8.0% x 2) + (9.0% x 1) + (7.75% x 1))/4).

The Initial Model Logic Problem

We evaluated the Initial Models for their efficacy in rewarding reasonableness and penalizing extremity using a variety of tests which are described in Chapter 3. Most of the Initial Models either provided weak incentives to "move to the middle" or, in a number of cases, incentives that operated in a contrary fashion – i.e., incentives that continued to encourage Parties to take extreme positions.

After considerable testing and analysis, we determined that the problem lay in the difficulty that the Initial Models had in distinguishing between Outcomes like:

Outcome A · Extreme/Extreme (Utilities)
Extreme/Immoderate (Intervenors)

Outcome B · Moderate/Moderate (Utilities)
Immoderate/Moderate (Intervenors)

In Outcome A, the Intervenors should be rewarded, because they adopted the only Immoderate Strategy in a sea of Extreme Strategies. Thus, the difference between the SPROR and the average recommendation (this difference being referred to hereafter as the SPROR Adjustment) should be *negative* (i.e., the SPROR should be less than the average, consistent with the interests of the Intervenors). But in Outcome B, the Intervenors should be penalized for adopting the only Immoderate Strategy given the Moderate Strategies adopted by the Utilities and the other Intervenor expert. Thus, the SPROR Adjustment in Outcome B should be *positive* (i.e., the SPROR should be greater than the average, consistent with the interests of the Utilities).

If, say, Model 2 is applied to Outcomes A and B in an Environment with a Reality Gap of 1% and a Strategy Gap of 0.25%, the resulting SPROR Adjustments are 0.52% and 0.40% respectively. Both of these adjustments are positive. Thus, Model 2 fails to distinguish between a situation where the "outlier" Strategy is more reasonable (i.e., Outcome A) and a situation where the "outlier" Strategy is more extreme (i.e., Outcome B).

We attempted to address the problem by incorporating Internal Difference adjustments (e.g., using Model 8 rather than Model 2). Sometimes this remedy improved one of the Outcomes but worsened the other; and, as in the example here, the incorporation of an Internal Difference adjustment sometimes fails to meaningfully change the final SPROR Adjustment at all. Curiously, if we *subtract* the Internal Difference rather than *adding* it to the raw differences, the perverse incentives are corrected for some of the Outcomes but are created in other Outcomes that had previously delivered satisfactory results.

The identification of the problem and these observations led us to consider that classic playground favourite - the "see-saw." The see-saw is a board that moves up and down across a fixed fulcrum point.

Is it possible to define a "fulcrum point" with *positive* Internal Difference adjustments on one side of the fulcrum and *negative* Internal Difference adjustments on the other side? (And, logically, a zero adjustment at the fulcrum point itself.) Our thinking about fulcrum points led us to develop what we refer to as a Weighted Average Fulcrum Adjustment (WAFA) Model. The WAFA Model resolves many of the concerns identified in the Initial Models.

Weighted Average Fulcrum Adjustment (WAFA) Models

After careful analysis, we concluded that the solution to the logic problem lay in establishing a "fulcrum point" (hereafter, a Fulcrum) based on the Internal Recommendation Spread of each Outcome.

At Internal Recommendation Spreads *below* the Fulcrum, the Internal Difference Adjustments are positive (i.e., the Internal Differences for both the Utilities and the Intervenors are *added* to the absolute values of the differences between each recommendation and the central value for all recommendations).[10] But at Internal Recommendation Spreads *above* the Fulcrum, the Internal Difference Adjustment is negative (i.e., the Internal Differences for both the Utilities and the Intervenors are *subtracted* from the absolute values of the differences between each recommendation and the central value for all recommendations). If the Internal Recommendation Spread associated with an Outcome equals the Fulcrum, then the Internal Difference adjustment is 0%.

The WAFA structure dramatically improves the efficacy of the Models and their ability to distinguish between illustrative situations like Outcomes A and B above.

Any of the Structures from the Initial Models can be modified to introduce an Internal Difference adjustment that changes from positive to negative at the Fulcrum. For our 24 WAFA Models, we chose the basic Structure which first appears in Model 4. In this Structure, the recommendation that is closest to the central value receives a weight of 4x. The recommendation that is furthest from the average receives a weight of 1x; and the

[10] We begin here to use the phrase "central value" rather than average, because some of the WAFA models use the midpoint rather than the average as the measure of central tendency.

remaining recommendations receive weights of 2x each. We chose this basic Structure, because it is the only Structure which ensures that *some* weight is given to each of the four recommendations.

The 24 WAFA Models reflect the various combinations of four specifications. First, we consider Fulcrums at 2.0%, 2.25% and 2.5%.[11]

Second, we calculate the difference between each recommendation and the central value for all recommendations using the average as the central value and, alternatively, the midpoint between the highest Utility recommendation and the lowest Intervenor recommendation as the central value.

Third, in the original Structure, the recommendation that is closest to the central value receives a weight of 4x. We also consider the possibility of giving this observation a weight of 6x.[12]

Finally, we consider two alternative methods for computing the weights that are attached to each recommendation. If there is a clearly-defined ranking of which recommendation is closest and which is furthest from the central value, then the two methods give the same result. If, however, there is a two- or three-way "tie" for the "closest" or the "furthest" recommendations, then it is unclear what weights should be assigned. These weighting methods are referred to as the Fixed Sum method and the Fixed Value method.

In the Fixed Sum method, we recalculate the weights so that each of the "tied" values receives the same weight subject to the sum of the weights remaining unchanged from 9 or 11.[13]

For example, suppose that the Utility recommendations are both 10.0%, the Intervenor recommendations are 7.0% and 7.25%, the central value is calculated using the average and the Fulcrum is set at 2.5%. The average of the four recommendations is 8.56%. The Internal Recommendation Spread is 2.75% (= 10.0% *less* 7.25%), which is greater than

---

[11] A Fulcrum of 2.25% is at the midpoint of the entire range of Recommendation Differences that we considered. The Strategic/Strategic Recommendation Difference in an Environment with 0% Reality Gap and 0.25% Strategy Gap is 0.5%. The Extreme/Extreme Recommendation Difference in an Environment with 0% Reality Gap and 0.5% Strategy Gap is 4.0%. The midpoint between 0.5% and 4.0% is 2.25%. To test the sensitivity of our results to other Fulcrums, we used alternative values 25 basis points above and below 2.25%.

[12] The significance of choosing a weight of 6x as an alternative is that 6x is the lowest weight consistent with the minimum difference recommendation having more than 50% of the overall weight from among the four recommendations (i.e., 6 *divided by* 11 > 50%). Thus, the WAFA Models with 6x weight give more weight to the minimum difference recommendation than is given to the other three recommendations combined. In contrast, the WAFA Models with 4x weight give slightly less weight to the minimum difference recommendation than is given to the other three recommendations combined (i.e., 4 *divided by* 9 < 50%).

[13] In those WAFA Models where the minimum difference recommendation receives a weight of 4x, the sum of the weights for the four recommendations is 9 (= 4 for the closest recommendation + 1 for the furthest recommendation + 2 for each of the other two recommendations). In those WAFA Models where the minimum difference recommendation receives a weight of 6x, the sum of the weights for the four recommendations is 11 (= 6 for the closest recommendation + 1 for the furthest recommendation + 2 for each of the other two recommendations).

the Fulcrum of 2.5%. Thus, the Internal Differences of 0% (= 10.0% *less* 10.0%) for the Utilities and 0.25% (= 7.25% *less* 7.0%) for the Intervenors are *subtracted* from the absolute value of the differences between each recommendation and the average. The calculated differences for each of the recommendations are 1.44% (= 10.0% *less* 8.56% *less* 0%), 1.44% (= 10.0% *less* 8.56% *less* 0%), 1.31% (= 8.56% *less* 7.0% *less* 0.25%) and 1.06% (= 8.56% *less* 7.25% *less* 0.25%). The 7.25% recommendation has the lowest calculated difference and is therefore given a weight of 4x in the determination of the SPROR. Both of the 10.0% values have calculated differences of 1.44%, which are the highest values. To keep the sum of the weights equal to 9, we assign a weight of 1.5 (= the average of 1 and 2) to both of the Utility recommendations. The remaining Intervenor observation (the 7.0%) receives a weight of 2. The weighted average SPROR is 8.11%; and the SPROR Adjustment is a *negative* 0.45% (= 8.11% *less* 8.56%).

In the Fixed Value method, we strictly adopt a weight of 4x or 6x for however many recommendations are associated with the lowest differences; and we adopt a weight of 1x for however many recommendations are associated with the highest differences. All differences between the highest and the lowest are assigned a weight of 2. The sum of the weights will vary depending on how many "ties" appear in each scenario; and the weighted average of the recommendations is computed by dividing the weighted recommendations by the sum of the weights associated with that Outcome.

There are 24 combinations of three Fulcrums (2.0%, 2.25% or 2.5%), two methods for determining weights (Fixed Sum and Fixed Value), two methods for determining the central value (Average and Midpoint) and two values for the minimum difference weight (4 or 6). The specifications for WAFA Models 25 – 48 are:

Model 25 · 2.0% Fulcrum · Fixed Sum Weights · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · Minimum Difference Weight = 4x

Model 26 · 2.25% Fulcrum · Fixed Sum Weights · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · Minimum Difference Weight = 4x

Model 27 · 2.5% Fulcrum · Fixed Sum Weights · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · Minimum Difference Weight = 4x

Model 28 · 2.0% Fulcrum · Fixed Value Weights · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · Minimum Difference Weight = 4x

Model 29 · 2.25% Fulcrum · Fixed Value Weights · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · Minimum Difference Weight = 4x

Model 30 · 2.5% Fulcrum · Fixed Value Weights · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · Minimum Difference Weight = 4x

Model 31 · 2.0% Fulcrum · Fixed Sum Weights · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · Minimum Difference Weight = 4x

Model 32 · 2.25% Fulcrum · Fixed Sum Weights · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · Minimum Difference Weight = 4x

Model 33 · 2.5% Fulcrum · Fixed Sum Weights · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · Minimum Difference Weight = 4x

Model 34 · 2.0% Fulcrum · Fixed Value Weights · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · Minimum Difference Weight = 4x

Model 35 · 2.25% Fulcrum · Fixed Value Weights · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · Minimum Difference Weight = 4x

Model 36 · 2.5% Fulcrum · Fixed Value Weights · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · Minimum Difference Weight = 4x

The specifications for Models 37 – 48 are the same as Models 25 – 36, except that the Minimum Difference Weights for Models 37-48 are 6x rather than 4x.

Consider the following examples that illustrate two of the WAFA Models.

Example 1. Assume that Model 29 is applied to a situation where the Utility witnesses recommend rates of return of 9.0% and 9.5%, and both of the Intervenor witnesses recommend 8.25%. The average recommendation is 8.75%; and the Internal Differences are 0.5% for the Utilities and 0% for the Intervenors. The Internal Recommendation Difference is 0.75% (= 9.0% *less* 8.25%), which is less than the 2.25% Fulcrum. Because the Internal Recommendation Difference is *less* than the Fulcrum, the Internal Differences are *added* in the calculation of the differences that will determine the weights for each recommendation. The calculated differences, adjusted for Internal Differences, are 1.25% (= 9.5% *less* 8.75% *plus* 0.5%), 0.75% (= 9.0% *less* 8.75% *plus* 0.5%), 0.5% (= 8.75% *less* 8.25% *plus* 0%) and 0.5% (= 8.75% *less* 8.25% *plus* 0%). Under the Fixed Value provisions of Model 29, the 9.5% recommendation receives a weight of 1x. The 9.0% recommendation receives a weight of 2x; and each of the 8.25% recommendations – being closest to the average as adjusted for Internal Differences – receives a weight of 4x. The weighted average SPROR is 8.50%, which is 25 basis points less than the average recommendation of 8.75%. Thus, the Intervenors are rewarded for their more moderate position and their smaller Internal Difference.

Example 2. Assume that Model 33 is applied to a situation where both of the Utility witnesses recommend rates of return of 9.75%, and the Intervenor witnesses recommend rates of return of 7.0% and 7.25%. The midpoint value is 8.38% (i.e., the midpoint between 9.75% and 7.0%). The Internal Recommendation Spread is 2.5%, which is exactly equal to the Fulcrum. Thus, there is no adjustment for Internal Differences. The differences between each recommendation and the midpoint value are 1.38% (= 9.75%

less 8.38%), 1.38% (= 9.75% *less* 8.38%), 1.38% (= 8.38% *less* 7.0%) and 1.13% (= 8.38% *less* 7.25%). Since Fixed Sum weights are used in Model 32, the sum of the weights must equal 9. The 7.25% recommendation has the lowest difference from the average and is therefore given a weight of 4x. The three remaining recommendations are "tied" for the furthest from the average at 1.38%. Therefore, each of these recommendations receives a weight of 1.67x (=(9-4) *divided by* 3) so that the sum of the weights equals 9 (= 4 + 1.67 + 1.67 + 1.67). The weighted average SPROR is 8.13%, which is 0.31% lower than the 8.44% average of the four recommendations.

The Arbitrary Weighting Concern and ICE Models

In Models 25 – 48, the recommendation closest to the central value is given a weight of 4x or 6x. The recommendation furthest from the central value is given a weight of 1x; and the other recommendations are given a weight of 2x. Although there is a logic to selecting 4x and 6x (i.e., the values that are just below and just above 50% of the total of the weights in each Fixed Sum case), we are concerned that there may remain a perception of arbitrariness in respect of how these weights were selected.

To address this potential concern, we designed Models 49 – 66, which use internally-calculated exponential (ICE) weights that are dictated by the absolute value of the differences between each recommendation and the central value and each recommendation's Internal Difference Adjustment.

To illustrate, suppose that the Utilities proffer recommendations of 10.0% and 10.0%; and the Intervenors sponsor witnesses whose recommendations are 7.0% and 7.25%. The average recommendation is 8.56%. The Internal Recommendation Spread is 2.75% (= 10.0% *less* 7.25%). If the Fulcrum is set at 2.25%, then the Internal Recommendation Spread exceeds the Fulcrum; and the Internal Differences are therefore *subtracted* from the raw differences between each recommendation and the average. The adjusted differences are 1.44% (= 10.0% *less* 8.56% *less* 0%), 1.44%, 1.31% (= 8.56% *less* 7.0% *less* 0.25%) and 1.06% (= 8.56% *less* 7.25% *less* 0.25%) respectively. As in all WAFA Models, the recommendation with the smallest adjusted difference (i.e., the 7.25%) should receive the highest weight; and the recommendations with the largest adjusted differences (i.e., the 10.0%) should receive the lowest weight.

In WAFA Models 25 – 48, the weights are based on exogenously-determined values such as 4, 2 and 1 or 6, 2 and 1. In ICE Models 49 – 66, however, the weights are calculated internally based on the ratio of the largest adjusted difference to each of the adjusted differences raised to the power of 1, 2 or 3. To illustrate, an ICE Model using an exponent of 2 applied to the data in the prior paragraph would assign a weight of 1.0 (i.e., 1.44% *divided by* 1.44% *raised to the power of* 2) to each of the 10.0% recommendations. A weight of 1.20 (i.e., 1.44% *divided by* 1.31% *raised to the power of* 2) would be assigned to the 7.0% recommendation; and a weight of 1.83 (i.e., 1.44% *divided by* 1.06% *raised to the power of* 2) would be assigned to the 7.25% recommendation.

The resulting weighted average SPROR is 8.28%; and the SPROR Adjustment is *negative* 0.28% (= 8.56% *less* 8.28%). Thus, the Intervenors are rewarded for their moderation.

The 18 ICE Models use alternative Fulcrums of 2.0%, 2.25% and 2.5%. The Average and the Midpoint are alternatively used to compute the central value; and the power to which the ratio of the adjusted differences is raised is alternatively 1, 2 or 3. The squaring (i.e., an exponent of 2) or cubing (i.e., an exponent of 3) of the ratios increases the difference between the weight assigned to the most central recommendation and the weight assigned to the furthest recommendation from the central value.

The specifications for ICE Models 49 – 66 are:

Model 49 · 2.0% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 1

Model 50 · 2.25% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 1

Model 51 · 2.5% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 1

Model 52 · 2.0% Fulcrum · Recommendation Differences Calculated *vis-à-vis* Midpoint Recommendation · ICE Exponent = 1

Model 53 · 2.25% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · ICE Exponent = 1

Model 54 · 2.5% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · ICE Exponent = 1

Model 55 · 2.0% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 2

Model 56 · 2.25% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 2

Model 57 · 2.5% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 2

Model 58 · 2.0% Fulcrum · Recommendation Differences Calculated *vis-à-vis* Midpoint Recommendation · ICE Exponent = 2

Model 59 · 2.25% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · ICE Exponent = 2

Model 60 · 2.5% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · ICE Exponent = 2

Model 61 · 2.0% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 3

Model 62 · 2.25% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 3

Model 63 · 2.5% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Average Recommendation · ICE Exponent = 3

Model 64 · 2.0% Fulcrum · Recommendation Differences Calculated *vis-à-vis* Midpoint Recommendation · ICE Exponent = 3

Model 65 · 2.25% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · ICE Exponent = 3

Model 66 · 2.5% Fulcrum · Recommendation Differences Calculated *vis-à-vis* the Midpoint Recommendation · ICE Exponent = 3

2.5 Winners and Losers

The objective is to identify Models which reward those Parties who do not take extreme positions and whose recommendations have modest Internal Differences. Conversely, the optimal Model should penalize those Parties who proffer extreme positions and whose recommendations have large Internal Differences. We have tested the efficacy of the 66 Models by applying each of them in 210 circumstances (= 70 Outcomes x 3 Environments).

The development of these tests is in two parts. First, we formally analyse the 70 Outcomes to determine which of the Parties in each circumstance *should* be rewarded (the Winners) or penalized (the Losers). Second, we establish four qualitative results for each test to determine if the appropriate Party is rewarded or penalized.

The terms “Win” and “Lose” have specific meanings, namely:

“Win” for the Utilities means an SPROR higher than the average of the four recommendations. “Win” for the Intervenors means an SPROR lower than the average of the four recommendations.

“Lose” for the Utilities means an SPROR lower than the average of the four recommendations. “Lose” for the Intervenors means an SPROR higher than the average of the four recommendations.

Establishing Winners and Losers for the 70 Outcomes

In this part of Chapter 2, we determine which of the Parties in each of the 70 Outcomes *should* be rewarded or penalized (i.e., which Party should be the Winner and which should be the Loser, where "Win" and "Lose" are defined above). The rules for establishing Winners and Losers are organized according to the Internal Differences of the Utility and Intervenor expert recommendations.

Internal Differences are expressed here using the number of Strategies that separate each Party's recommendations. Thus, if a Party's recommendations are Moderate/Moderate, then that Party's Internal Difference is 0. Alternatively, if a Party's recommendations are Immoderate/Strategic, then that Party's Internal Difference is 2 (Immoderate to Moderate and Moderate to Strategic).

An Internal Difference of, say, 0/1 means that the Utility experts have the same recommendation (i.e., the same Strategy); and the Intervenor experts have recommendations that differ by one Strategy "notch" (i.e., either 25 basis points or 50 basis points, depending on the Strategy Gap in the test Environment).

Table 2.3 shows possible combinations of Internal Differences and the rules that apply in each circumstance.

Table 2.3

Internal Differences and Winner/Loser Rules

Utility Internal Differences \ Intervenor Internal Differences	0	1	2	3
0	A	C	E	B
1	C	A	D	B
2	E	D	A	B
3	B	B	B	A

Rule A. If the Internal Differences are 0/0, 1/1, 2/2 or 3/3, then the results for all Models will be a simple average of the recommendations. These scenarios are not part of the tests for effectiveness of the 66 Models, because all of the Models produce this result.

Rule B. Any Party who proffers a Strategy Combination where the Internal Difference is 3 (i.e., Extreme/Strategic) should Lose.[14] For example, if the Utility recommendations are Extreme/Strategic and the Intervenor recommendations are Immoderate/Moderate, then the Utilities Lose. This rule may appear harsh, because it implies that an Extreme/ Strategic position would lose to an Extreme/Extreme position. Nevertheless, each Party has the reasonable ability to avoid filing evidence that could be considered Extreme/ Strategic (i.e., widely-disparate recommendations). Therefore, the risk of being caught under Rule B is manageable and within the reasonable control of the Parties.

Rule C. If the Internal Difference is 0/1 or 1/0, then the Party whose recommendations are closest to Forecast should Win. For example, if the Utility recommendations are Extreme/Immoderate and the Intervenor recommendations are Extreme/Extreme, then the Utilities should Win; and the Intervenors should Lose. Alternatively, if the Utilities are Immoderate/Moderate and the Intervenors are Moderate/Moderate, then the Intervenors should Win; and the Utilities should Lose.

Rule D. If the Internal Difference is 1/2 or 2/1, then there must be at least one Strategy in common between the Utilities and the Intervenors. The Winner is determined by which of the other recommendations is closer to Forecast. For example, if the Utilities are Extreme/Immoderate and the Intervenors are Immoderate/Strategic, then the Strategy in common is Immoderate. Of the remaining Strategies, the Strategic Strategy of the Intervenors is closer to Forecast than the Extreme Strategy of the Utilities. Therefore, the Intervenors should Win; and the Utilities should Lose. Alternatively, if the Utilities are Moderate/Strategic and the Intervenors are Immoderate/Strategic, then the Strategy in common is Strategic. Of the remaining Strategies, the Immoderate Strategy of the Intervenors is further away from Forecast than the Moderate Strategy of the Utilities. Therefore, the Utilities should Win; and the Intervenors should Lose.

Rule E. If the Internal Difference is 0/2 or 2/0, then :

1. Subject to (2) below, the Party whose recommendations are closest to Forecast should Win. For example, if the Utility recommendations are Extreme/Strategic and the Intervenor recommendations are Strategic/ Strategic, then the Intervenors should Win; and the Utilities should Lose. Alternatively, if the Utility recommendations are Immoderate/ Strategic and the Intervenor recommendations are Immoderate/ Immoderate, then the Utilities should Win; and the Intervenors should Lose.

[14] The assumption is made throughout that neither the Utilities nor the Intervenors will proffer a Forecast recommendation – i.e., a recommendation equal to what they truly believe the Commission will actually award. Thus, the widest Internal Difference is 3, arising from an Extreme/Strategic Strategy Combination.

2. With a 0/2 or 2/0 Internal Difference, it is possible that the Party with the Internal Difference of 2 might be sponsoring recommendations which reflect a Strategy Combination that "encircles" the Strategies implicit in the recommendations of the Party with the Internal Difference of 0. If so, then the Party with the 0 Internal Difference Wins. For example, suppose that the Utility recommendations are Extreme/Moderate; and the Intervenor recommendations are Immoderate/Immoderate. The Intervenor recommendations are "encircled" by the Extreme/Moderate recommendations of the Utilities. In recognition of the objective of minimizing Internal Differences, the Intervenors would Win; and the Utilities would Lose. Alternatively, if the Utilities are Moderate/Moderate and the Intervenors are Immoderate/Strategic, then the Utilities Win; and the Intervenors Lose.

Table 2.4 shows each of the 70 Outcomes with the Strategy Combination that should Win shown in ***bold italic.***

Table 2.4

WINNERS AND LOSERS FOR 70 OUTCOMES

Outcomes	Party 1 Strategy Combinations	Party 2 Strategy Combinations	Internal Differences	Applicable Rule
1 and 36	Extreme/Extreme	***Extreme/Immoderate***	0/1	C
2 and 37	Extreme/Extreme	***Extreme/Moderate***	0/2	E
3 and 38	***Extreme/Extreme***	Extreme/Strategic	0/3	B
4 and 39	Extreme/Extreme	***Immoderate/Moderate***	0/1	C
5 and 40	Extreme/Extreme	***Immoderate/Strategic***	0/2	E
6 and 41	Extreme/Extreme	***Moderate/Strategic***	0/1	C
7 and 42	Extreme/Immoderate	***Extreme/Moderate***	1/2	D
8 and 43	***Extreme/Immoderate***	Extreme/Strategic	1/3	B
9 and 44	Extreme/Immoderate	***Immoderate/Immoderate***	1/0	C
10 and 45	Extreme/Immoderate	***Immoderate/Strategic***	1/2	D
11 and 46	Extreme/Immoderate	***Moderate/Moderate***	1/0	C
12 and 47	Extreme/Immoderate	***Strategic/Strategic***	1/0	C
13 and 48	***Extreme/Moderate***	Extreme/Strategic	2/3	B
14 and 49	Extreme/Moderate	***Immoderate/Immoderate***	2/0	E
15 and 50	Extreme/Moderate	***Immoderate/Moderate***	2/1	D
16 and 51	Extreme/Moderate	***Moderate/Moderate***	2/0	E
17 and 52	Extreme/Moderate	***Moderate/Strategic***	2/1	D
18 and 53	Extreme/Moderate	***Strategic/Strategic***	2/0	E
19 and 54	Extreme/Strategic	***Immoderate/Immoderate***	3/0	B
20 and 55	Extreme/Strategic	***Immoderate/Moderate***	3/1	B
21 and 56	Extreme/Strategic	***Immoderate/Strategic***	3/2	B
22 and 57	Extreme/Strategic	***Moderate/Moderate***	3/0	B
23 and 58	Extreme/Strategic	***Moderate/Strategic***	3/1	B
24 and 59	Extreme/Strategic	***Strategic/Strategic***	3/0	B
25 and 60	Immoderate/Immoderate	***Immoderate/Moderate***	0/1	C
26 and 61	Immoderate/Immoderate	***Immoderate/Strategic***	0/2	E
27 and 62	Immoderate/Immoderate	***Moderate/Strategic***	0/1	C
28 and 63	Immoderate/Moderate	***Immoderate/Strategic***	1/2	D
29 and 64	Immoderate/Moderate	***Moderate/Moderate***	1/0	C
30 and 65	Immoderate/Moderate	***Strategic/Strategic***	1/0	C
31 and 66	Immoderate/Strategic	***Moderate/Moderate***	2/0	E
32 and 67	Immoderate/Strategic	***Moderate/Strategic***	2/1	D
33 and 68	Immoderate/Strategic	***Strategic/Strategic***	2/0	E
34 and 69	Moderate/Moderate	***Moderate/Strategic***	0/1	C
35 and 70	Moderate/Strategic	***Strategic/Strategic***	1/0	C

Note: The Strategy Combinations in ***bold italic*** are considered to be the Winners - i.e., the moderate Strategy Combinations that we hope the optimal Model will encourage. Outcomes 1-35 assume that the Utilities are Party 1, and the Intervenors are Party 2. Outcomes 36-70 assume that the Intervenors are Party 1, and the Utilities are Party 2.

Establishing Four Qualitative Results

Each of the 13,860 tests (= 70 Outcomes x 66 Models x 3 Environments) is assigned a qualitative result that describes its ability to provide the appropriate incentives. The four results are:

1. a Simple Average Outcome;

2. a Positive Incentive in which moderation is meaningfully rewarded and extremity is penalized;

3. a Negative Incentive in which extremity is meaningfully rewarded and moderation is penalized; or

4. a Close Outcome in which the rewards and penalties are too weak to be considered meaningful.

A Simple Average Outcome arises when the Simple Average Rule applies, and the SPROR is therefore calculated as the simple average of the lowest Utility and highest Intervenor recommendations.

A Positive Incentive occurs when the Party that *should* be the Winner and the party that *should* be the Loser as shown in Table 2.4 actually turn out to be the Winner and the Loser respectively – i.e., when the incentives encourage moderation and discourage extremity. For example, if the Utility recommendations are Extreme/Immoderate and the Intervenor recommendations are Immoderate/Strategic (Outcome 10), then the result consistent with a Positive Incentive is that the Intervenors will Win; and the Utilities will Lose. In other words, the Intervenors will be rewarded for taking a relatively more moderate position; and the Utilities will be penalized for taking a relatively more extreme position.

In this example, if the calculated SPROR is *less* than the average of the four recommendations, then the Intervenors are the Winners; and the Utilities are the Losers. The test is then said to result in a Positive Incentive, because the test results are consistent with the incentives that we wish to encourage (i.e., moderation rather than extremity). If, however, the calculated SPROR is *greater* than the average of the four recommendations, then the Intervenors are the Losers; and the Utilities are the Winners. The test would then be said to result in a Negative Incentive, because the test results are contrary to the incentives that we wish to encourage.

A Close Outcome means that the difference between the SPROR and the average of the four recommendations is less than 10 basis points – i.e., the rewards and penalties are too weak to be considered meaningful.

2.6 Criteria for Analyzing Test Results

Before presenting the test results, we set out set out six criteria and related statistics that are used to determine the efficacy of each Model. These criteria are: Fairness, an Incentive to Reduce the Recommendation Spread, an Incentive to Minimise Internal Differences, a Reasonable SPROR Adjustment[15], Incentive Strength and a Reasonable Ratio of Positive to Negative Incentives.

Fairness

Fairness means that a symmetric end result will occur irrespective of whether the positions are taken by Utilities or Intervenors. To illustrate, if the Utilities adopt strategies that are Immoderate/Moderate and if the Intervenors adopt strategies that are Moderate/Moderate and if the result is a *downward* adjustment to the simple average recommendation of, say, 50 basis points, then a 50 basis points *upward* adjustment to the simple average recommendation should result from the Utilities having adopted a Moderate/Moderate strategy and the Intervenors having adopted an Immoderate/ Moderate strategy.

Each of the Models has been tested for fairness. None of the Models yield asymmetric results. Thus, none of the Models are "unfair."

Incentive to Reduce the Recommendation Spread

We test each Model's ability to encourage the Parties to reduce the Recommendation Spread by reference to two statistics.

First, we compute the average SPROR Adjustment associated with a Party's electing the Extreme/Extreme Strategy Combination. If the Utilities adopt the Extreme/Extreme Strategy Combination, then the expected SPROR Adjustment should be negative. If the Intervenors adopt the Extreme/Extreme Strategy Combination, then the expected SPROR Adjustment should be positive.

Second, in a similar fashion, we determine the average SPROR Adjustment associated with a particular Party's electing each of the ten Strategy Combinations (i.e., Extreme/Extreme, Extreme/Immoderate…Strategic/Strategic).[16] We then create three group averages – an average for the four Strategy Combinations which have at least one Extreme recommendation (i.e., Extreme/Extreme, Extreme/Immoderate, Extreme/ Moderate and Extreme/Strategic), an average for the three Strategy Combinations which have Immoderate but not Extreme as one of the recommendations (Immoderate/

[15] The difference between each Model's calculated SPROR and the simple average of the four recommendations is referred to as the SPROR Adjustment.
[16] The discussion which follows is from the perspective of the Utilities; however, the same discussion and conclusions apply to the Intervenors, because the Model results are symmetric.

Immoderate, Immoderate/Moderate and Immoderate/Strategic) and an average for the three remaining Strategy Combinations that have neither an Extreme nor an Immoderate recommendation (i.e., Moderate/Moderate, Moderate/Strategic and Strategic/Strategic).

For a Model with the correct incentives, the average SPROR Adjustment for Utility Moderate and Strategic Strategies will be greater than the average SPROR Adjustment for Utility Immoderate Strategies which, in turn, will be greater than the average SPROR Adjustment for Utility Extreme Strategies. Stated another way, the Utilities should expect the largest *positive* SPROR Adjustment from adopting Strategies that fall into the Moderate or Strategic group; and they should expect the largest *negative* SPROR Adjustment from adopting Strategies that fall into the Extreme group.[17]

Incentive to Minimise Internal Differences

Each of the ten Strategy Combinations is associated with an Internal Difference from 0 (e.g., Moderate/Moderate) to 3 (e.g., Extreme/Strategic). Ideally, a model should encourage Parties to minimise the Internal Difference between the recommendations of their experts. To test for the latter, we calculated the average SPROR Adjustment for Strategy Combinations with Internal Differences of 0 (i.e., Extreme/Extreme, Immoderate/Immoderate, Moderate/Moderate and Strategic/Strategic), Strategy Combinations with Internal Differences of 1 (i.e., Extreme/Immoderate, Immoderate/Moderate and Moderate/Strategic) and Strategy Combinations with Internal Differences of 2 or 3 (i.e., Extreme/Moderate, Extreme/Strategic and Immoderate/Strategic).

For a Model with the correct incentives, the average SPROR Adjustment for Utility Internal Differences of 0 should be greater than the average SPROR Adjustment for Utility Internal Differences of 1 which, in turn, should be greater than the average SPROR Adjustment for Utility Internal Differences of 2 or 3. Stated another way, the Utilities should expect their largest *positive* SPROR Adjustment from adopting Strategies where Internal Differences are minimised.

Reasonable SPROR Adjustment

The SPROR Adjustment is a reflection of the incentive that Parties have to act reasonably. Thus, the calculated SPROR Adjustment must be strong enough to provide a meaningful incentive to act reasonably but should not be so strong that it produces results which are likely to be unacceptable to the Commission or the Parties.

[17] The reverse is true for Intervenors. The Intervenors should expect their largest *negative* SPROR Adjustment from adopting Strategies that fall into the Moderate or Strategic group; and they should expect the largest *positive* SPROR Adjustment from adopting Strategies that fall into the Extreme group.

Chart 1.2 displays differences between Utility and Intervenor recommended rates of return for the period 1984 – 2017. From 1984 to 2000, differences were in the range of roughly 2.0 – 2.5%; and the re-establishment of recommendation differences within this range is a worthy objective given that these differences have generally exceeded 3.0% for some while.

The Commission and its predecessors have never to our knowledge awarded rates of return entirely outside the range of the expert recommendations in a proceeding. Thus, the maximum SPROR Adjustment that the Commission would have made during this early period which we hope to re-establish would be 1.25% (i.e., 2.50% *divided by* 2).

As a result, we consider that SPROR Adjustments over 125 basis points are likely to produce results which will be unacceptable to the Commission or the Parties.[18] For each Model, we calculate the maximum SPROR Adjustment across all 70 Outcomes. We also calculate the maximum SPROR Adjustment across Outcomes 25-35 and 60-70. The latter group are those Outcomes which do not include Extreme Strategies for either Party. If a Model's maximum SPROR Adjustment is greater than 100 basis points, then we regard that Model as having potentially unhelpful results.[19]

Incentive Strength

Aside from the magnitude of the SPROR Adjustment, incentive strength is measured by two statistics – namely, that proportion of the 210 tests (= 70 Outcomes x 3 Environments) for each Model which result in Positive Incentives or Negative Incentives and the proportion of tests which result in Positive Incentive SPROR Adjustments. As a matter of judgment, we consider that no less than 75% of tests should result in either a Positive or a Negative Incentive; and we conclude that no less than 50% of tests should result in a Positive Incentive.

The former statistic is referred to as the Percent of Outcomes Other than Average (POOA) and is calculated by dividing the sum of the number of tests yielding Positive Incentives and Negative Incentives by the total number of tests. The latter statistic is referred to as Positive Incentives as a Percent of Total Outcomes (PIPTO) and is calculated by dividing the sum of the number of tests yielding Positive Incentives by the total number of tests.

[18] To illustrate, suppose that the simple average of the four recommendations is, say, 8.5% and the range of recommendations is 7.0-10.0%. A Utility would likely find a rate of return less than 7.25% (= 8.5% *less* 1.25%) to be wholly unacceptable and probably outside the bounds of reasonableness when judged by the decisions of regulators in other jurisdictions. Similarly, an Intervenor would likely find a rate of return greater than 9.75% (= 8.5% *plus* 1.25%) to be wholly unacceptable for the same reason.

[19] The 100 basis points is based on a Recommendation Spread of 2.0% (i.e., at the lower end of the approximate 2.0 – 2.5% range from Chart 1.2 for 1984 - 2000).

Ratio of Positive to Negative Incentives

Different Outcomes in different Environments produce an array of positive and negative incentives. Recall that a "positive incentive" is one which rewards the more reasonable Party and penalizes the more extreme Party. A "negative incentive" has the reverse effect. Clearly, then, positive incentives are preferred to negative incentives. We consider that – at the least - the number of Outcomes with a Positive Incentive should meaningfully exceed the number of Outcomes with a Negative Incentive such that the ratio of positive to negative is no less than 150%.

2.7 Test Procedures

We calculated the SPROR and the SPROR Adjustments for each of the 66 Models using the 70 Outcomes and the three Environments. Thus, each Model was applied in 210 circumstances; and the results in each circumstance were assigned to one of the four categories discussed in Part 2.5 above.

Next, the data underlying the criteria described in Part 2.6 were compiled. Models which satisfied all the criteria were retained for more detailed analysis.

The remaining Models were examined with a view to selecting one or two Models with incentives that will encourage the Parties to seek a middle ground and a range of Outcomes that would be considered reasonable by the Commission and the Parties.

Chapter 3

DETERMINING THE OPTIMAL GAME

In Chapter 3, we apply the tests and evaluation criteria discussed in Chapter 2 to determine rules for the optimal game or Model. Chapter 3 is divided into the following parts.

3.1 Compiling the Summary of Results

SPROR Adjustments are calculated for each of the 66 Models under 210 different circumstances (= 70 Outcomes x 3 Environments). The 70 Outcomes are the 70 combinations of Strategies that may be adopted by the Utilities and the Intervenors.[1] The three Environments contemplate: (i) a Reality Gap of 0% and a Strategy Gap of 0.25%; (ii) a Reality Gap of 1.0% and a Strategy Gap of 0.25%; and (iii) a Reality Gap of 0% and a Strategy Gap of 0.5%.

A Summary of Results table is prepared for each of the 66 Models. The Summary of Results sets out the underlying assumptions of each Model and aggregated statistics for the 210 separate tests. Table 3.1 on the next page shows the Summary of Results for Model 41 and is illustrative.

[11] The 70 Outcomes exclude situations where both the Utilities and the Intervenors adopt the same Strategies or have the same combination of Strategies.

Table 3.1

SUMMARY OF RESULTS · MODEL 41

ID Adjustment Fulcrum Value	2.25%	**All Outcomes/Environments**	
ID % Adjustment > Fulcrum Value	(100.00%)	Max Adjustment	1.03%
ID % Adjustment < Fulcrum Value	100.00%	Min Adjustment	0.19%
Fixed Sum or Fixed Value Weights	FV	**Outcomes 25-35 and 61-70**	
Average or Midpoint	Avg	Max Adjustment	0.78%
Minimum Difference Weight	6	Min Adjustment	0.18%

	Reality Gap = 0% Strategy Gap of 0.25%	Reality Gap = 1% Strategy Gap of 0.25%	Reality Gap = 0% Strategy Gap of 0.5%	All Outcomes
Simple Average Outcomes	0	0	0	0
Positive Incentives	44	50	50	144
Negative Incentives	24	18	18	60
Close Outcomes	2	2	2	6
Total Outcomes	70	70	70	210
Recommendation Spreads	2.00%	3.00%	4.00%	

Fairness · No Asymmetries

Asymmetries	0	0	0	0	Yes

Incentive to Reduce Recommendation Spread · Average SPROR Adjustment for Extreme/Extreme is <0% · Average SPROR Adjustment Should Decline with More Extreme Strategy Groups

Average SPROR Adjustment for Extreme/Extreme Strategy				(0.16%)	Yes
Average SPROR Adjustment by Strategy Groups	Mod/Strategic 0.14%	Immoderate (0.05%)	Extreme (0.14%)		Yes

Incentive to Minimise Internal Differences · Average SPROR Adjustment Should Decline with Higher Internal Differences

Average SPROR Adjustment by Internal Differences	ID = 0 0.10%	ID = 1 0.04%	ID = 2 or 3 (0.26%)		Yes

Reasonable SPROR Adjustment · Maximum Absolute Value Adjustment

Maximum Absolute Value of SPROR Adjustment for All Outcomes <= 1.25%	1.03%	Yes
Maximum Absolute Value of SPROR Adjustment for Outcomes 25-35 and 61-70 <= 1.0%	0.78%	Yes

Incentive Strength · POOA >= 75% and PIPTO >= 50%

Percent of Outcomes Other than Average (POOA)	97.1%	Yes
Positive Incentives as a Percent of Total Outcomes (PIPTO)	68.6%	Yes

Ratio of Positive to Negative Incentives · PIPTO/NIPTO >= 150%

Ratio of Positive Incentives to Negative Incentives (PIPTO/NIPTO)	240%	Yes

The information at the top of the Summary of Results describes the Model and its characteristics. The Maximum and Minimum SPROR Adjustments for all Outcomes and Environments are displayed alongside the Maximum and Minimum SPROR Adjustments for Outcomes 25-35 and 60-70. The significance of these Outcomes is that these are the Outcomes which have *no* Extreme Strategies. The Maximum and Minimum All Outcomes/Environments are used to test whether the range of SPROR Adjustments are reasonable.

The next part of the Summary of Results shows the number of Simple Average Outcomes, Positive Incentives, Negative Incentives and Close Outcomes sorted by Environment. Of the 210 circumstances under which Model 41 is applied, 144 result in Positive Incentives, 60 result in Negative Incentives and 6 result in a Close Outcome (i.e., an SPROR Adjustment less than 10 basis points).

Beneath the table are six areas which provide the nine sets of data used to test each of the criteria described in Part 2.6. In addition to the raw data, "Yes" or "No" appears in the right-most column to indicate whether or not the Model satisfies each criterion.

3.2 Applying the Evaluation Criteria

Table 3.2 on the next three pages shows the numerical values of the statistics compiled on each Model's *Summary of Results*.

Table 3.2

DATA FOR APPLICATION OF EVALUATION CRITERIA

Model	Description	Asymm	Adjustment for Ext/Ext	Average SPROR Adjustments						Maximum SPROR Adjust		POOA	PIPTO	PIPTO/NIPTO
				Mod/Strat	Immoderate	Extreme	ID = 0	ID = 1	ID = 2 or 3	All Outcomes	25-35/60-70			
1	SPROR Is Recommendation Associated with Minimum Difference from Average No Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	(0.56%)	(0.10%)	0.05%	0.14%	(0.37%)	0.23%	0.39%	1.63%	1.13%	71%	30%	70%
2	SPROR Is Average Recommendation Excluding Maximum Difference from Average No Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.28%	0.10%	(0.03%)	(0.13%)	0.22%	(0.08%)	(0.32%)	0.79%	0.58%	71%	42%	142%
3	Min Difference Party Excludes Max Difference Recom of Other Party No Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	(0.23%)	(0.04%)	0.02%	0.07%	(0.15%)	0.07%	0.20%	0.63%	0.46%	71%	30%	70%
4	SPROR Is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others No Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	(0.03%)	0.02%	(0.00%)	(0.02%)	(0.00%)	0.03%	(0.03%)	0.12%	0.00%	6%	5%	500%
5	SPROR Is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others No Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.07%	0.05%	(0.01%)	(0.06%)	0.07%	0.00%	(0.14%)	0.38%	0.25%	54%	34%	171%
6	SPROR Is Avg Recom Excluding Those with Differences Outside One Std Dev No Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.70%	0.19%	(0.07%)	(0.26%)	0.51%	(0.23%)	(0.68%)	1.88%	1.38%	71%	42%	142%
7	SPROR Is Recommendation Associated with Minimum Difference from Average Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.70%	0.19%	(0.07%)	(0.25%)	0.51%	(0.26%)	(0.64%)	1.88%	1.38%	71%	42%	142%
8	SPROR Is Average Recommendation Excluding Maximum Difference from Average Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.28%	0.10%	(0.03%)	(0.13%)	0.22%	(0.08%)	(0.32%)	0.79%	0.58%	71%	42%	142%
9	Min Difference Party Excludes Max Difference Recom of Other Party Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.28%	0.10%	(0.03%)	(0.13%)	0.22%	(0.08%)	(0.32%)	0.79%	0.58%	71%	42%	142%
10	SPROR Is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.25%	0.08%	(0.03%)	(0.10%)	0.19%	(0.08%)	(0.26%)	0.65%	0.49%	71%	42%	142%
11	SPROR Is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.39%	0.12%	(0.04%)	(0.16%)	0.29%	(0.12%)	(0.40%)	1.00%	0.75%	71%	42%	142%
12	SPROR Is Avg Recom Excluding Those with Differences Outside One Std Dev Adjustment for Internal Differences · Rule 1 Average Limit = 1.0%	0	0.57%	0.19%	(0.06%)	(0.28%)	0.44%	(0.14%)	(0.69%)	2.29%	1.58%	71%	42%	142%
13	SPROR Is Recommendation Associated with Minimum Difference from Average No Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	(0.56%)	(0.15%)	0.07%	0.19%	(0.42%)	0.23%	0.49%	1.63%	1.13%	99%	34%	53%
14	SPROR Is Average Recommendation Excluding Maximum Difference from Average No Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.28%	0.12%	(0.04%)	(0.15%)	0.24%	(0.07%)	(0.37%)	0.79%	0.58%	100%	66%	192%
15	Min Difference Party Excludes Max Difference Recom of Other Party No Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	(0.23%)	(0.09%)	0.03%	0.11%	(0.19%)	0.06%	0.28%	0.63%	0.46%	100%	34%	52%
16	SPROR Is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others No Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	(0.03%)	0.01%	0.00%	(0.01%)	(0.02%)	0.03%	(0.01%)	0.15%	0.00%	9%	8%	800%
17	SPROR Is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others No Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.07%	0.05%	(0.01%)	(0.07%)	0.07%	0.01%	(0.15%)	0.38%	0.25%	77%	54%	238%
18	SPROR Is Avg Recom Excluding Those with Differences Outside One Std Dev No Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.70%	0.25%	(0.09%)	(0.32%)	0.56%	(0.22%)	(0.79%)	1.88%	1.38%	100%	66%	192%

Model	Description	Asymm	Adjustment for Ext/Ext	Average SPROR Adjustments						Maximum SPROR Adjust		POOA	PIPTO	PIPTO/NIPTO
				Mod/Strat	Immoderate	Extreme	ID = 0	ID = 1	ID = 2 or 3	All Outcomes	25-35/60-70			
19	SPROR Is Recommendation Associated with Minimum Difference from Average Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.70%	0.24%	(0.09%)	(0.29%)	0.56%	(0.26%)	(0.74%)	1.88%	1.38%	100%	66%	192%
20	SPROR Is Average Recommendation Excluding Maximum Difference from Average Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.28%	0.12%	(0.04%)	(0.15%)	0.24%	(0.07%)	(0.37%)	0.79%	0.58%	100%	66%	192%
21	Min Difference Party Excludes Max Difference Recom of Other Party Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.28%	0.12%	(0.04%)	(0.15%)	0.24%	(0.07%)	(0.37%)	0.79%	0.58%	100%	66%	192%
22	SPROR Is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.25%	0.09%	(0.03%)	(0.12%)	0.20%	(0.08%)	(0.29%)	0.65%	0.49%	100%	66%	192%
23	SPROR Is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.39%	0.15%	(0.05%)	(0.19%)	0.32%	(0.12%)	(0.46%)	1.00%	0.75%	100%	66%	192%
24	SPROR Is Avg Recom Excluding Those with Differences Outside One Std Dev Adjustment for Internal Differences · Rule 1 Average Limit = 0%	0	0.57%	0.29%	(0.09%)	(0.37%)	0.52%	(0.12%)	(0.86%)	2.29%	1.58%	100%	66%	192%
25	WAFA · 2.0% Fulcrum · FS Weights · Average · Minimum Difference Weight = 4	0	(0.10%)	0.07%	(0.03%)	(0.06%)	0.01%	0.06%	(0.10%)	0.57%	0.40%	81%	58%	254%
26	WAFA · 2.25% Fulcrum · FS Weights · Average · Minimum Difference Weight = 4	0	(0.08%)	0.09%	(0.03%)	(0.09%)	0.06%	0.04%	(0.19%)	0.57%	0.47%	90%	65%	252%
27	WAFA · 2.5% Fulcrum · FS Weights · Average · Minimum Difference Weight = 4	0	(0.00%)	0.09%	(0.03%)	(0.11%)	0.11%	0.00%	(0.24%)	0.57%	0.47%	90%	63%	236%
28	WAFA · 2.0% Fulcrum · FV Weights · Average · Minimum Difference Weight = 4	0	(0.17%)	0.09%	(0.03%)	(0.06%)	(0.00%)	0.07%	(0.10%)	0.88%	0.63%	81%	58%	254%
29	WAFA · 2.25% Fulcrum · FV Weights · Average · Minimum Difference Weight = 4	0	(0.14%)	0.11%	(0.04%)	(0.11%)	0.07%	0.04%	(0.21%)	0.88%	0.63%	90%	65%	252%
30	WAFA · 2.5% Fulcrum · FV Weights · Average · Minimum Difference Weight = 4	0	(0.02%)	0.11%	(0.04%)	(0.13%)	0.14%	(0.02%)	(0.27%)	0.88%	0.61%	90%	63%	236%
31	WAFA · 2.0% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 4	0	(0.12%)	0.04%	(0.01%)	(0.02%)	(0.04%)	0.06%	(0.01%)	0.57%	0.40%	95%	62%	186%
32	WAFA · 2.25% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 4	0	(0.12%)	0.06%	(0.02%)	(0.06%)	0.01%	0.04%	(0.09%)	0.57%	0.40%	95%	67%	233%
33	WAFA · 2.5% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 4	0	(0.06%)	0.06%	(0.02%)	(0.07%)	0.05%	0.02%	(0.13%)	0.57%	0.32%	95%	66%	223%
34	WAFA · 2.0% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 4	0	(0.20%)	0.04%	(0.02%)	(0.01%)	(0.07%)	0.10%	0.01%	0.88%	0.63%	94%	62%	191%
35	WAFA · 2.25% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 4	0	(0.16%)	0.14%	(0.05%)	(0.14%)	0.10%	0.04%	(0.26%)	1.03%	0.78%	97%	69%	240%
36	WAFA · 2.5% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 4	0	(0.12%)	0.07%	(0.03%)	(0.07%)	0.04%	0.03%	(0.14%)	0.88%	0.48%	94%	66%	230%
37	WAFA · 2.0% Fulcrum · FS Weights · Average · Minimum Difference Weight = 6	0	(0.14%)	0.08%	(0.03%)	(0.06%)	(0.00%)	0.07%	(0.10%)	0.76%	0.53%	96%	64%	200%
38	WAFA · 2.25% Fulcrum · FS Weights · Average · Minimum Difference Weight = 6	0	(0.12%)	0.12%	(0.04%)	(0.12%)	0.08%	0.05%	(0.23%)	0.76%	0.61%	97%	69%	240%
39	WAFA · 2.5% Fulcrum · FS Weights · Average · Minimum Difference Weight = 6	0	(0.03%)	0.12%	(0.04%)	(0.13%)	0.13%	0.01%	(0.29%)	0.76%	0.61%	97%	68%	229%
40	WAFA · 2.0% Fulcrum · FV Weights · Average · Minimum Difference Weight = 6	0	(0.19%)	0.10%	(0.04%)	(0.07%)	(0.01%)	0.08%	(0.11%)	1.03%	0.73%	96%	64%	200%

Model	Description	Asymm	Adjustment for Ext/Ext	Average SPROR Adjustments						Maximum SPROR Adjust		POOA	PIPTO	PIPTO/NIPTO
				Mod/Strat	Immoderate	Extreme	ID = 0	ID = 1	ID = 2 or 3	All Outcomes	25-35/60-70			
41	WAFA · 2.25% Fulcrum · FV Weights · Average · Minimum Difference Weight = 6	0	(0.16%)	0.14%	(0.05%)	(0.14%)	0.10%	0.04%	(0.26%)	1.03%	0.78%	97%	69%	240%
42	WAFA · 2.5% Fulcrum · FV Weights · Average · Minimum Difference Weight = 6	0	(0.03%)	0.14%	(0.06%)	(0.16%)	0.18%	(0.02%)	(0.34%)	1.03%	0.78%	97%	68%	229%
43	WAFA · 2.0% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 6	0	(0.17%)	0.04%	(0.01%)	(0.01%)	(0.06%)	0.08%	0.02%	0.76%	0.53%	93%	61%	188%
44	WAFA · 2.25% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 6	0	(0.16%)	0.07%	(0.02%)	(0.06%)	0.00%	0.05%	(0.09%)	0.76%	0.53%	93%	66%	238%
45	WAFA · 2.5% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 6	0	(0.10%)	0.07%	(0.02%)	(0.07%)	0.04%	0.03%	(0.13%)	0.76%	0.44%	93%	65%	227%
46	WAFA · 2.0% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 6	0	(0.23%)	0.04%	(0.02%)	(0.00%)	(0.09%)	0.11%	0.04%	1.03%	0.73%	91%	60%	191%
47	WAFA · 2.25% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 6	0	(0.23%)	0.08%	(0.03%)	(0.06%)	(0.01%)	0.07%	(0.09%)	1.03%	0.73%	91%	65%	243%
48	WAFA · 2.5% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 6	0	(0.15%)	0.08%	(0.03%)	(0.08%)	0.04%	0.03%	(0.14%)	1.03%	0.63%	91%	64%	231%
49	ICE · 2.0% Fulcrum · Average · ICE Exponent = 1	0	(0.07%)	0.05%	(0.02%)	(0.04%)	0.01%	0.02%	(0.05%)	0.89%	0.35%	68%	50%	294%
50	ICE · 2.25% Fulcrum · Average · ICE Exponent = 1	0	(0.05%)	0.06%	(0.02%)	(0.06%)	0.04%	0.01%	(0.11%)	0.89%	0.37%	76%	57%	300%
51	ICE · 2.5% Fulcrum · Average · ICE Exponent = 1	0	(0.02%)	0.06%	(0.02%)	(0.07%)	0.07%	0.00%	(0.14%)	0.81%	0.37%	72%	53%	280%
52	ICE · 2.0% Fulcrum · Midpoint · ICE Exponent = 1	0	(0.17%)	0.02%	(0.01%)	0.01%	(0.07%)	0.05%	0.08%	2.00%	0.47%	53%	34%	185%
53	ICE · 2.25% Fulcrum · Midpoint · ICE Exponent = 1	0	(0.16%)	0.03%	(0.01%)	(0.02%)	(0.04%)	0.04%	0.02%	2.00%	0.47%	46%	34%	288%
54	ICE · 2.5% Fulcrum · Midpoint · ICE Exponent = 1	0	(0.07%)	0.03%	(0.01%)	(0.03%)	0.01%	0.02%	(0.06%)	0.75%	0.19%	44%	34%	343%
55	ICE · 2.0% Fulcrum · Average · ICE Exponent = 1	0	(0.16%)	0.08%	(0.03%)	(0.05%)	(0.00%)	0.04%	(0.06%)	1.17%	0.61%	85%	60%	242%
56	ICE · 2.25% Fulcrum · Average · ICE Exponent = 2	0	(0.13%)	0.10%	(0.03%)	(0.10%)	0.06%	0.03%	(0.18%)	1.17%	0.65%	93%	68%	263%
57	ICE · 2.5% Fulcrum · Average · ICE Exponent = 2	0	(0.03%)	0.10%	(0.03%)	(0.12%)	0.11%	0.00%	(0.24%)	1.17%	0.65%	91%	65%	243%
58	ICE · 2.0% Fulcrum · Midpoint · ICE Exponent = 2	0	(0.25%)	0.03%	(0.01%)	0.01%	(0.10%)	0.09%	0.08%	1.74%	0.83%	79%	52%	196%
59	ICE · 2.25% Fulcrum · Midpoint · ICE Exponent = 2	0	(0.23%)	0.06%	(0.02%)	(0.04%)	(0.04%)	0.07%	(0.04%)	1.74%	0.83%	77%	57%	286%
60	ICE · 2.5% Fulcrum · Midpoint · ICE Exponent = 2	0	(0.13%)	0.06%	(0.02%)	(0.06%)	0.02%	0.04%	(0.10%)	1.08%	0.33%	76%	56%	281%
61	ICE · 2.0% Fulcrum · Average · ICE Exponent = 3	0	(0.19%)	0.11%	(0.04%)	(0.07%)	(0.01%)	0.06%	(0.08%)	1.24%	0.84%	90%	62%	217%
62	ICE · 2.25% Fulcrum · Average · ICE Exponent = 3	0	(0.15%)	0.13%	(0.04%)	(0.13%)	0.08%	0.05%	(0.24%)	1.24%	0.84%	94%	68%	254%
63	ICE · 2.5% Fulcrum · Average · ICE Exponent = 3	0	(0.04%)	0.14%	(0.05%)	(0.15%)	0.15%	0.00%	(0.31%)	1.24%	0.84%	94%	67%	241%
64	ICE · 2.0% Fulcrum · Midpoint · ICE Exponent = 3	0	(0.33%)	0.04%	(0.02%)	0.01%	(0.13%)	0.11%	0.11%	2.44%	1.02%	91%	59%	182%
65	ICE · 2.25% Fulcrum · Midpoint · ICE Exponent = 3	0	(0.30%)	0.07%	(0.02%)	(0.05%)	(0.05%)	0.09%	(0.05%)	2.44%	1.02%	91%	65%	243%
66	ICE · 2.5% Fulcrum · Midpoint · ICE Exponent = 3	0	(0.17%)	0.07%	(0.02%)	(0.07%)	0.02%	0.05%	(0.12%)	1.35%	0.43%	91%	64%	231%

The evaluation criteria are then applied to the data to determine whether a Model satisfies each criterion. Table 3.3 below and the on the next two pages is entitled *Application of Evaluation Criteria*. The "Yes" and "No" which appear for each of the nine data items are taken from the right-most column of the Summary of Results pages for each Model.

Table 3.3

APPLICATION OF EVALUATION CRITERIA

Model	Description	Fairness	Reduce Recom Spread		Minimise	Reasonable SPROR Adjust		Incentive Strength		Pos to Neg	Criteria
			Ext/Ext	Adj by Group	Int Diff	All Out	25-35/60-70	POOA	PIPTO	Incentive	Satisfied
1	SPROR Is Recommendation Associated with Minimum Difference from Average No Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	Yes	No	No	No	No	No	No	No	2
2	SPROR Is Average Recommendation Excluding Maximum Difference from Average No Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	Yes	Yes	No	No	No	5
3	Min Difference Party Excludes Max Difference Recom of Other Party No Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	Yes	No	No	Yes	Yes	No	No	No	4
4	SPROR Is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others No Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes	6
5	SPROR Is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others No Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes	6
6	SPROR Is Avg Recom Excluding Those with Differences Outside One Std Dev No Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	No	No	No	No	No	3
7	SPROR Is Recommendation Associated with Minimum Difference from Average Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	No	No	No	No	No	3
8	SPROR Is Average Recommendation Excluding Maximum Difference from Average Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	Yes	Yes	No	No	No	5
9	Min Difference Party Excludes Max Difference Recom of Other Party Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	Yes	Yes	No	No	No	5
10	SPROR Is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	Yes	Yes	No	No	No	5
11	SPROR Is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	Yes	Yes	No	No	No	5
12	SPROR Is Avg Recom Excluding Those with Differences Outside One Std Dev Adjustment for Internal Differences - Rule 1 Average Limit = 1.0%	Yes	No	Yes	Yes	No	No	No	No	No	3
13	SPROR Is Recommendation Associated with Minimum Difference from Average No Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	Yes	No	No	No	No	Yes	No	No	3
14	SPROR Is Average Recommendation Excluding Maximum Difference from Average No Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	8
15	Min Difference Party Excludes Max Difference Recom of Other Party No Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	Yes	No	No	Yes	Yes	Yes	No	No	5
16	SPROR Is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others No Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes	6
17	SPROR Is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others No Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	8
18	SPROR Is Avg Recom Excluding Those with Differences Outside One Std Dev No Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	6

Model	Description	Fairness	Reduce Recom Spread		Minimise Int Diff	Reasonable SPROR Adjust		Incentive Strength		Pos to Neg Incentive	Criteria Satisfied
			Ext/Ext	Adj by Group		All Out	25-35/60-70	POOA	PIPTO		
19	SPROR is Recommendation Associated with Minimum Difference from Average Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	6
20	SPROR is Average Recommendation Excluding Maximum Difference from Average Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	8
21	Min Difference Party Excludes Max Difference Recom of Other Party Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	8
22	SPROR is 4x Min Difference Recom, 1x Max Difference Recom and 2x Others Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	8
23	SPROR is 2x Min Difference Recom, 0x Max Difference Recom and 1x Others Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	8
24	SPROR is Avg Recom Excluding Those with Differences Outside One Std Dev Adjustment for Internal Differences - Rule 1 Average Limit = 0%	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	6
25	WAFA - 2.0% Fulcrum - FS Weights - Average - Minimum Difference Weight = 4	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
26	WAFA - 2.25% Fulcrum - FS Weights - Average - Minimum Difference Weight = 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
27	WAFA - 2.5% Fulcrum - FS Weights - Average - Minimum Difference Weight = 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
28	WAFA - 2.0% Fulcrum - FV Weights - Average - Minimum Difference Weight = 4	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
29	WAFA - 2.25% Fulcrum - FV Weights - Average - Minimum Difference Weight = 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
30	WAFA - 2.5% Fulcrum - FV Weights - Average - Minimum Difference Weight = 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
31	WAFA - 2.0% Fulcrum - FS Weights - Midpoint - Minimum Difference Weight = 4	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
32	WAFA - 2.25% Fulcrum - FS Weights - Midpoint - Minimum Difference Weight = 4	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
33	WAFA - 2.5% Fulcrum - FS Weights - Midpoint - Minimum Difference Weight = 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
34	WAFA - 2.0% Fulcrum - FV Weights - Midpoint - Minimum Difference Weight = 4	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	7
35	WAFA - 2.25% Fulcrum - FV Weights - Midpoint - Minimum Difference Weight = 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
36	WAFA - 2.5% Fulcrum - FV Weights - Midpoint - Minimum Difference Weight = 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
37	WAFA - 2.0% Fulcrum - FS Weights - Average - Minimum Difference Weight = 6	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
38	WAFA - 2.25% Fulcrum - FS Weights - Average - Minimum Difference Weight = 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
39	WAFA - 2.5% Fulcrum - FS Weights - Average - Minimum Difference Weight = 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
40	WAFA - 2.0% Fulcrum - FV Weights - Average - Minimum Difference Weight = 6	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8

Model	Description	Fairness	Reduce Recom Spread		Minimise	Reasonable SPROR Adjust		Incentive Strength		Pos to Neg	Criteria
			Ext/Ext	Adj by Group	Int Diff	All Out	25-35/60-70	POOA	PIPTO	Incentive	Satisfied
41	WAFA - 2.25% Fulcrum - FV Weights - Average - Minimum Difference Weight = 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
42	WAFA - 2.5% Fulcrum - FV Weights - Average - Minimum Difference Weight = 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
43	WAFA - 2.0% Fulcrum - FS Weights - Midpoint - Minimum Difference Weight = 6	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	7
44	WAFA - 2.25% Fulcrum - FS Weights - Midpoint - Minimum Difference Weight = 6	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
45	WAFA - 2.5% Fulcrum - FS Weights - Midpoint - Minimum Difference Weight = 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
46	WAFA - 2.0% Fulcrum - FV Weights - Midpoint - Minimum Difference Weight = 6	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	7
47	WAFA - 2.25% Fulcrum - FV Weights - Midpoint - Minimum Difference Weight = 6	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
48	WAFA - 2.5% Fulcrum - FV Weights - Midpoint - Minimum Difference Weight = 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
49	ICE - 2.0% Fulcrum - Average - ICE Exponent = 1	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	7
50	ICE - 2.25% Fulcrum - Average - ICE Exponent = 1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
51	ICE - 2.5% Fulcrum - Average - ICE Exponent = 1	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	8
52	ICE - 2.0% Fulcrum - Midpoint - ICE Exponent = 1	Yes	Yes	No	No	No	Yes	No	No	Yes	4
53	ICE - 2.25% Fulcrum - Midpoint - ICE Exponent = 1	Yes	Yes	Yes	No	No	Yes	No	No	Yes	5
54	ICE - 2.5% Fulcrum - Midpoint - ICE Exponent = 1	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes	6
55	ICE - 2.0% Fulcrum - Average - ICE Exponent = 1	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
56	ICE - 2.25% Fulcrum - Average - ICE Exponent = 2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
57	ICE - 2.5% Fulcrum - Average - ICE Exponent = 2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
58	ICE - 2.0% Fulcrum - Midpoint - ICE Exponent = 2	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	6
59	ICE - 2.25% Fulcrum - Midpoint - ICE Exponent = 2	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	7
60	ICE - 2.5% Fulcrum - Midpoint - ICE Exponent = 2	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
61	ICE - 2.0% Fulcrum - Average - ICE Exponent = 3	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	8
62	ICE - 2.25% Fulcrum - Average - ICE Exponent = 3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
63	ICE - 2.5% Fulcrum - Average - ICE Exponent = 3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9
64	ICE - 2.0% Fulcrum - Midpoint - ICE Exponent = 3	Yes	Yes	No	No	No	No	Yes	Yes	Yes	5
65	ICE - 2.25% Fulcrum - Midpoint - ICE Exponent = 3	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	6
66	ICE - 2.5% Fulcrum - Midpoint - ICE Exponent = 3	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	7

The number of "Yes" entries for each Model is counted and appears in the final column of Table 3.3. Models 26, 27, 29, 30, 33, 35, 36, 38, 39, 41, 42, 45, 48, 50, 56, 57, 62 and 63 achieve a score of 9 out of 9 and are retained for further analysis.

3.3 Excluding Models with Weak Incentives

A principal concern is that whatever Model we recommend should have sufficiently strong incentives to encourage moderation and discourage extremity. Two of the evaluation criteria that we described in Part 2.6 are specifically designed to measure the incentive to reduce the Recommendation Spread – namely, the negative SPROR Adjustment associated with a Party's electing the Extreme/Extreme Strategy Combination and the "steepness" of the SPROR Adjustments calculated as the difference between the SPROR Adjustments associated with the most desirable Outcomes (i.e., the Moderate/Strategic Strategies) and the least desirable Outcomes (i.e., the Extreme Strategies).

Table 3.4 presents the Extreme/Extreme SPROR Adjustment values and the "Steepness" values (i.e., the differences between the average Moderate/Strategic SPROR Adjustment and the average Extreme SPROR Adjustment) for the 18 Models identified in Part 3.2.

Table 3.4

ELIMINATION OF MODELS BASED ON WEAK INCENTIVES

Model	Description	Adjustment for Ext/Ext	Average SPROR Adjustments				Weak Incentives
			Mod/Strat	Immoderate	Extreme	"Steepness"	
26	WAFA · 2.25% Fulcrum · FS Weights · Average · Minimum Difference Weight = 4	(0.08%)	0.09%	(0.03%)	(0.09%)	0.19%	X
27	WAFA · 2.5% Fulcrum · FS Weights · Average · Minimum Difference Weight = 4	(0.00%)	0.09%	(0.03%)	(0.11%)	0.20%	X
29	WAFA · 2.25% Fulcrum · FV Weights · Average · Minimum Difference Weight = 4	(0.14%)	0.11%	(0.04%)	(0.11%)	0.22%	
30	WAFA · 2.5% Fulcrum · FV Weights · Average · Minimum Difference Weight = 4	(0.02%)	0.11%	(0.04%)	(0.13%)	0.24%	X
33	WAFA · 2.5% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 4	(0.06%)	0.06%	(0.02%)	(0.07%)	0.13%	X
35	WAFA · 2.25% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 4	(0.19%)	0.07%	(0.02%)	(0.06%)	0.13%	X
36	WAFA · 2.5% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 4	(0.12%)	0.07%	(0.03%)	(0.07%)	0.14%	X
38	WAFA · 2.25% Fulcrum · FS Weights · Average · Minimum Difference Weight = 6	(0.12%)	0.12%	(0.04%)	(0.12%)	0.24%	
39	WAFA · 2.5% Fulcrum · FS Weights · Average · Minimum Difference Weight = 6	(0.03%)	0.12%	(0.04%)	(0.13%)	0.25%	X
41	WAFA · 2.25% Fulcrum · FV Weights · Average · Minimum Difference Weight = 6	(0.16%)	0.14%	(0.05%)	(0.14%)	0.28%	
42	WAFA · 2.5% Fulcrum · FV Weights · Average · Minimum Difference Weight = 6	(0.03%)	0.14%	(0.06%)	(0.16%)	0.30%	X
45	WAFA · 2.5% Fulcrum · FS Weights · Midpoint · Minimum Difference Weight = 6	(0.10%)	0.07%	(0.02%)	(0.07%)	0.14%	X
48	WAFA · 2.5% Fulcrum · FV Weights · Midpoint · Minimum Difference Weight = 6	(0.15%)	0.08%	(0.03%)	(0.08%)	0.16%	X
50	ICE · 2.25% Fulcrum · Average · ICE Exponent = 1	(0.05%)	0.06%	(0.02%)	(0.06%)	0.12%	X
56	ICE · 2.25% Fulcrum · Average · ICE Exponent = 2	(0.13%)	0.10%	(0.03%)	(0.10%)	0.21%	
57	ICE · 2.5% Fulcrum · Average · ICE Exponent = 2	(0.03%)	0.10%	(0.03%)	(0.12%)	0.22%	X
62	ICE · 2.25% Fulcrum · Average · ICE Exponent = 3	(0.15%)	0.13%	(0.04%)	(0.13%)	0.27%	
63	ICE · 2.5% Fulcrum · Average · ICE Exponent = 3	(0.04%)	0.14%	(0.05%)	(0.15%)	0.29%	X

Note: Models with an Average SPROR Adjustment of greater than *minus* 0.1% for Extreme/Extreme strategies are eliminated; and Models with less than 0.2% difference between the Average SPROR Adjustment for Moderate/Strategic Outcomes and Extreme Outcomes are also eliminated.

We eliminated from further consideration Models having average SPROR Adjustments for Extreme/Extreme positions of greater than *minus* 0.1%. Stated otherwise, we excluded Models where the *penalty* for adopting an Extreme/Extreme Strategy was less than 10 basis points.

We also excluded from further consideration those Models where the difference between the average SPROR Adjustment for Moderate/Strategic Strategies and Extreme Strategies was less than 20 basis points.

An "X" appears in the final column of Table 3.4 for those Models having weak incentives on either of the two criteria; and these Models are eliminated from further consideration. Models 29, 38, 41, 56 and 62 remain for analysis.

3.4 Analyzing the Final Selection Models

We approached the final selection and the development of our recommendations based on a detailed review of the quantitative data shown in Table 3.5 on the following page. For each data series, we indicate the "first and second place" models based solely on that data series and display the "first place" data items in ***bold italic*** and the "second place" data items in **bold**. To illustrate, the "first and second place" Models based on the largest negative SPROR Adjustment for Extreme/Extreme strategies are Models 41 and 62, with SPROR Adjustments of (0.16%) and (0.15%) respectively. Thus, the ***(0.16%)*** appears in bold italic in the Adjustment for Ext/Ext column of the table; and the **(0.15%)** appears in bold.

The "Steepness" of Average SPROR Adjustments are calculated as the differences between the SPROR Adjustments associated with the most desirable Outcomes (i.e., the Moderate/Strategic Strategies and the Internal Difference = 0 Strategies) and the least desirable Outcomes (i.e., the Extreme Strategies and the Internal Difference = 2 or 3 Strategies). The larger the difference, the more "steep" is the downward slope of the line which rewards moderation and penalizes extremity. Thus, the larger the difference, the greater the incentive provided by the Model.

Those Models with meaningful but more modest Minimum SPROR Adjustments are considered superior to those with more extreme Minimum SPROR Adjustments. Those Models with higher Percent Outcomes Other than Average (POOAs), higher Positive Incentives as a Percent of Total Outcomes (PIPTOs) and higher ratios of Positive Incentives to Negative Incentives (PIPTO/NIPTOs) are superior to Models with lower values on these statistics.

Table 3.5

ANALYSIS OF FINAL SELECTION MODELS

Model	Description	Adjustment for Ext/Ext	"Steepness" of Average SPROR Adjustments		Min SPROR Adjustments		POOA	PIPTO	PIPTO/ NIPTO	Final Ranking
			Strategy	ID	All	25-35/60-70				
29	2.25% Fulcrum · FV · Average Minimum Difference Wgt = 4	(0.14%)	0.22%	0.28%	**0.88%**	**0.63%**	90%	65%	252%	4
38	2.25% Fulcrum · FS · Average Minimum Difference Wgt = 6	(0.12%)	0.24%	0.31%	***0.76%***	***0.61%***	***97%***	***69%***	240%	2/3
41	2.25% Fulcrum · FV · Average Minimum Difference Wgt = 6	***(0.16%)***	***0.28%***	***0.36%***	1.03%	0.78%	***97%***	***69%***	240%	1
56	2.25% Fulcrum · Average ICE Exponent = 2	(0.13%)	0.21%	0.24%	1.17%	0.65%	93%	68%	***263%***	5
62	2.25% Fulcrum · Average ICE Exponent = 3	**(0.15%)**	**0.27%**	**0.32%**	1.24%	0.84%	94%	68%	**254%**	2/3

Note: Entries in ***bold italic*** denote a "first place" ranking. Entries in **bold** denote a "second place" ranking.

We assign Model 41 a Final Ranking of 1 in the right-most column of Table 3.4 for three reasons. First, Model 41 has more "first place" rankings based on the evaluation criteria than any of the other Models. Second, Model 41 has the strongest incentives to reduce the Recommendation Spread (i.e., the largest negative SPROR Adjustment for Extreme/ Extreme Strategies and the "steepest" average SPROR Adjustments as a Party moves from Moderate/Strategic to Extreme). Third, the calculation of the Fixed Value weights in the SPROR calculation is more straightforward than the calculation of the Fixed Sum weights in Model 38 or the ICE Weights in Models 56 and 62. Thus, we assign Model 41 a Final Ranking of 1 in the right-most column of Table 3.4. Model 41 is a WAFA Model with a 2.25% Fulcrum, the central value calculated as the Average, Fixed Value weights and a weight for the Minimum Difference Recommendation of 6x.

The considerations respecting second and third place rankings are mixed. After Model 41, Model 38 has the second largest number of "first place" rankings. That fact makes Model 38 an obvious candidate for a Final Ranking of 2. However, Model 62 has slightly superior incentives to reduce the Recommendation Spread based on the two evaluation criteria discussed in the previous paragraph. Moreover, if the Commission were concerned about the assignment of fixed, pre-assigned weights (Model 41) and preferred to have the weights calculated internally (i.e., an ICE Model), then Model 62 could be assigned the Final Ranking of 2. On balance, we are unable to sufficiently distinguish between the desirability of Model 38 and that of Model 62. Therefore, we assign each a ranking of "2/3."

We assign the remaining Models 29 and 56 rankings of 4 and 5 respectively based largely on the slightly stronger incentives to reduce Recommendation Spreads for Model 29.

3.5 Recommendations

We conclude that Model 41 is the optimal game that should be adopted by the Commission to provide Parties with appropriate and fair incentives to moderate their positions. The formal rules of Model 41 are:

1. Calculate the simple average of the four recommendations.

2. Calculate the absolute values of the differences between each recommendation and the average.

3. For Outcomes with Internal Difference Recommendations *less* than the 2.25% Fulcrum, *add* the Internal Difference to the differences calculated in (2). For Outcomes with Internal Difference Recommendations *greater* than the 2.25% Fulcrum, *subtract* the Internal Difference from the differences calculated in (2). If the Outcome's Internal Difference Recommendation is exactly equal to 2.25%, make no adjustment to the values from (2).

4. Calculate a weighted average of the recommendations where those recommendations having the minimum difference from the values calculated in (3) receive a weight of 6. Those recommendations having the maximum difference from the values calculated in (3) receive a weight of 1; and the other recommendations each have a weight of 2. The SPROR is the weighted average of the four recommendations.[2]

Consider the following sample calculation using Model 41. Assume that the Utilities proffer recommendations of 10.0% and 9.25%; and the Intervenors support recommendations of 7.25% and 7.75%.[3] The simple average of the four recommendations is 8.56%; and the absolute values of the differences between each recommendation and the average are 1.44%, 0.69%, 1.31% and 0.81%. The Internal Difference for the Utilities is 75 basis points (= 10.0% *less* 9.25%); and the Internal Difference for the Intervenors is 50 basis points (= 7.75% *less* 7.25%). The Internal Recommendation Spread is 1.5% (= 9.25% *less* 7.75%) and is therefore less than the 2.25% Fulcrum. Thus, the Internal Differences are *added* to the absolute values of the differences. The results are 2.19%, 1.44%, 1.81% and 1.31%.

The 7.75% recommendation is associated with the lowest difference of 1.31%. As a result, the 7.75% receives a weight of 6x in the weighted average calculation. The 10.0% recommendation is associated with the highest difference of 2.19%. Thus, the 10.0% receives a weight of 1x in the weighted average calculation. The other recommendations – the 9.25% and the 7.25% - each receive a weight of 2x. The weighted average Starting Point Rate of Return is therefore 8.14% (= ((7.75% x 6) + (9.25% x 2) + (7.25% x 2) + (10.0% x 1)) *divided by* 11). The 8.14% represents an SPROR Adjustment of *negative* 42 basis points (= 8.14% *less* 8.56%).

The Intervenors emerge as the Winners in this Model, because: (i) the Utilities proffered recommendations with a larger Internal Difference than the Intervenors; and (ii) excluding the Strategic recommendations from both the Utilities and the Intervenors, the remaining Strategy for the Intervenors (i.e., Immoderate) is more moderate than the remaining Strategy for the Utilities (i.e., Extreme).

Although we prefer Model 41, we would not be averse to using either Model 38 or Model 62 if the Commission and the Parties preferred one of those alternatives.

---

[2] Model 41 uses Fixed Value weights. It is therefore possible to have more than one recommendation assigned a weight of 6x or 1x. The SPROR is the sum of the weighted recommendations *divided by* the sum of the weights.

[3] These values are found in an Environment with a Reality Gap of 1.0% and a Strategy Gap of 0.25%. In this Environment, the Utilities are Extreme/Strategic; and the Intervenors are Immoderate/Strategic (i.e., Outcome 21).

3.6 An Observation Regarding the Importance of Simultaneous Filings

In the present regulatory regime, Utilities typically file applications; and Intervenors respond. But in a generic proceeding, there is no formal applicant and therefore no reason why the Parties cannot file simultaneously and thus secure the advantages of a closed bid auction which presses Parties to a middle ground.

The fairness of using, say, Model 41 as the basis for determining the Starting Point Rate of Return depends on simultaneous filing – i.e., a "closed bid" arrangement. If, alternatively, Intervenors were aware of Utility recommendations in advance, then the Intervenors would always be able to "beat" the Utilities and create a zero or negative SPROR Adjustment. Moreover, Intervenors would have no incentive to moderate their own recommendations – they would simply pick the combination of recommendations that produced the lowest SPROR Adjustment given their *a priori* knowledge of the choices made by the Utilities.

Table 3.6 on the next page shows the impact of permitting one Party (in this case, Intervenors) to "go second" in the process.

Table 3.6

SPROR ADJUSTMENTS FROM MODEL 41 AVERAGED ACROSS THREE ENVIRONMENTS

	Intervenors										Intervenors' Optimal Choice
	Ext/Ext	Ext/Immod	Ext/Mod	Ext/Strat	Immod/Immod	Immod/Mod	Immod/Strat	Mod/Mod	Mod/Strat	Strat/Strat	
Utilities											
Ext/Ext	0.00%	(0.42%)	(0.34%)	(0.12%)	0.00%	(0.37%)	(0.15%)	0.00%	(0.19%)	0.00%	(0.42%)
Ext/Immod	0.42%	0.00%	(0.16%)	0.52%	0.37%	0.00%	0.46%	0.19%	0.00%	(0.56%)	(0.56%)
Ext/Mod	0.34%	0.16%	0.00%	0.42%	0.15%	(0.46%)	0.00%	(0.62%)	(0.39%)	(0.52%)	(0.62%)
Ext/Strat	0.12%	(0.52%)	(0.42%)	0.00%	(0.68%)	(0.44%)	(0.34%)	(0.58%)	(0.36%)	(0.48%)	(0.68%)
Immod/Immod	0.00%	(0.37%)	(0.15%)	0.68%	0.00%	(0.19%)	0.62%	0.00%	0.56%	0.00%	(0.37%)
Immod/Mod	0.37%	0.00%	0.46%	0.44%	0.19%	0.00%	0.39%	(0.56%)	0.00%	(0.46%)	(0.56%)
Immod/Strat	0.15%	(0.46%)	0.00%	0.34%	(0.62%)	(0.39%)	0.00%	(0.52%)	(0.31%)	(0.42%)	(0.62%)
Mod/Mod	0.00%	(0.19%)	0.62%	0.58%	0.00%	0.56%	0.52%	0.00%	0.46%	0.00%	(0.19%)
Mod/Strat	0.19%	0.00%	0.39%	0.36%	(0.56%)	0.00%	0.31%	(0.46%)	0.00%	(0.36%)	(0.56%)
Strat/Strat	0.00%	0.56%	0.52%	0.48%	0.00%	0.46%	0.42%	0.00%	0.36%	0.00%	0.00%
										Average	(0.46%)

For example, if the Intervenors know that the Utilities are taking an Extreme/Extreme position (i.e., the first row of Table 3.6), then the Intervenors will take an Extreme/Immoderate position, because that combination of Strategies produces the largest *downward* adjustment to the SPROR (i.e., *minus* 42 basis points).

If the Intervenors know that the Utilities are taking an Immoderate/Moderate position (i.e., the sixth row of Table 3.6), then the Intervenors will logically take a Moderate/Moderate position, because that combination of Strategies produces the largest downward adjustment to the SPROR (i.e., *minus* 56 basis points). In both cases, the Intervenors have an incentive to squeeze in just below the level of extremity reflected in the Utility recommendations. The Intervenors have no incentive to move any further to the middle, because their interests lie in maximizing the downward adjustment to the SPROR.[4]

The right-most column of Table 3.6 shows the SPROR Adjustments that would logically result from permitting the Intervenors to have foreknowledge of the positions to be taken by the Utilities. On average, this knowledge gives the Intervenors an advantage of 46 basis points. Interestingly – and perhaps coincidentally – this theoretically-produced value is somewhat consistent with the apparent "advantage" that Intervenors have enjoyed in recent years when the average recommendation is compared to the regulatory decision. To illustrate, for decisions handed down between 2011 and 2017, the difference between the average recommendation and the regulatory decision is 37 basis points.

[4] This same analysis would apply to the Utilities if the Utilities received the advantage of filing second. The only difference is that the Utilities would receive a positive SPROR Adjustment, whereas the Intervenors enjoy a negative SPROR Adjustment.

Chapter 4

DETERMINING THE OPTIMAL STRATEGY

4.1 Introduction

In Chapter 3, we developed a recommended Model (Model 41) that encouraged the Utilities and Intervenors to "move to the middle" in their rate of return recommendations. In Chapter 4, we determine the optimal strategy for the Parties assuming the implementation of Model 41. If the recommended Model does its job properly, then the optimal strategy should encourage the Parties to take positions that are closer to Forecast and farther from Extreme.

It may be obvious that the optimal strategy depends on what assumptions are made about how the other Party will behave and what Strategies the other Party will select. But to place all Parties on a level playing field and secure the benefits of a closed bid auction, simultaneous disclosure is required. Therefore, the possibility that the other Party will adopt a particular Strategy is not a certainty but a probability; and the array of possible Strategies must therefore be overlaid with a probability Distribution in order to determine the *a priori* optimal strategy.

For purposes of simplification, the discussion in Chapter 4 considers only the optimal strategy decision from the perspective of the Utilities who do not know what Strategies will be adopted by the Intervenors. The mirror-image of this problem (i.e., from the perspective of the Intervenors) yields the same results but with a change in the signs of most numbers.

In Part 4.2, we develop alternative probability Distributions for Strategy Combinations that might logically be adopted by the Intervenors. For each of the 10 Strategy Combinations that might be adopted by the Utilities (e.g., Extreme/Extreme, Extreme/ Immoderate, Extreme/Moderate… Strategic/Strategic), a weighted average SPROR Adjustment is calculated.[1] The individual SPROR Adjustments are computed by combining each Utility Strategy Combination with each of the possible Intervenor Strategy Combinations; and the weights are the probabilities from the assumed Distribution. The weighted average SPROR Adjustments are then ranked from highest to lowest to determine the optimal strategy for the Utilities.

Equation 4.1 shows the formal method of calculation.

[1] Mathematically: (4!/ 2! 2!) + 4 = 10 for each of the Utilities and the Intervenors. As with the analysis in Chapters 2 and 3, we assume that neither the Utilities nor the Intervenors will proffer a Forecast recommendation.

Equation 4.1 $$SA_{i,j} = \sum_{k=1}^{10} (SA_{i,k} \times p_{j,k}) \text{ subject to } \sum_{k=1}^{10} p_{j,k} = 1.0 \text{ for all } j$$

where:

$SA_{i,j}$ = Expected SPROR Adjustment for Utility Strategy i and Intervenor Strategy Combination Probability Distribution j

$SA_{j,k}$ = SPROR Adjustment for Utility Strategy Combination i and Intervenor Strategy Combination k

$p_{j,k}$ = Probability of Intervenors Adopting Strategy Combination k under Intervenor Probability Distribution j

Chapter 4 is divided into the following parts.

4.2 Consideration of Alternative Probability Distributions

In Part 4.2, we develop five alternative probability Distributions which describe the likelihood of each of the ten possible Strategy Combinations that could be adopted by the Intervenors.[2] The exact Distribution is unknown; however, we have considered alternatives with a reasonable underlying logic and which conform to the following rules.

[2] The ten possibilities are Extreme/Extreme, Extreme/Immoderate, Extreme/Moderate, Extreme/ Strategic, Immoderate/Immoderate, Immoderate/Moderate, Immoderate/Strategic, Moderate/Moderate, Moderate/Strategic and Strategic/Strategic.

1. No probability should exceed 15%.[3]

2. No probability should be less than 5%.[4]

3. Meaningful differences should exist between those Strategies which are considered to be more likely and those which are considered to be less likely.[5]

Distribution 1

Probability Distribution 1 assumes that each of the ten Intervenor Strategy Combinations has equal probability. Thus, the probability assigned to each Strategy Combination is 10% (= 100% *divided by* 10). The sum of these probabilities is 100%.

Distribution 2

Probability Distribution 2 assumes that the highest probability is attached to the three Intervenor Strategy Combinations with an Internal Difference of 1 (i.e., Extreme/ Immoderate, Immoderate/Moderate and Moderate/Strategic). The next highest probability is attached to the four Intervenor Strategy Combinations with an Internal Difference of 0 (i.e., Extreme/Extreme, Immoderate/Immoderate, Moderate/Moderate and Strategic/Strategic); and the lowest probability is attached to the three Strategy Combinations with Internal Differences of 2 or 3 (i.e., Extreme/Moderate, Extreme/Strategic and Immoderate/Strategic).

The basis for this ranking of probabilities is that the Parties are aware that the method for determining the SPROR rewards those with smaller Internal Differences and penalizes those with larger Internal Differences, all else equal. Therefore, although some Internal Difference may be desirable, that difference is unlikely to be a 2 or 3, which would imply rate of return differences of 50-150 basis points based on Strategy Gaps of 0.25% and 0.50%.

Consistent with the rules for assigning probabilities and the ranking described here, we assigned a 15% weight to each of the three Strategy Combinations with an Internal Difference of 1. We assigned a 10% weight to each of the four Strategy Combinations with an Internal Difference of 0 and a 5% weight to each of the three Strategy Combinations with an Internal Difference of 2 or 3. The sum of these probabilities is 100% (= (15% x 3) + (10% x 4) + (5% x 3)).

---

[3] This rule is relaxed in the circumstances of Distribution 4, because three of the ten Outcomes are assigned a probability of 20%.
[4] This rule is relaxed in the circumstances of Distribution 4, because four of the ten Outcomes are assigned a probability of 0%.
[5] This rule is relaxed in the circumstances of Distribution 1, because all of the ten Outcomes are purposely assigned the same probability.

Distribution 3

Probability Distribution 3 assumes that the highest probability is attached to the four Intervenor Strategy Combinations with an Internal Difference of 0 (i.e., Extreme/Extreme, Immoderate/Immoderate, Moderate/Moderate and Strategic/ Strategic). The next highest probability is attached to the three Strategy Combinations with an Internal Difference of 1 (i.e., Extreme/Immoderate, Immoderate/Moderate and Moderate/Strategic); and the lowest probability is attached to the three Strategy Combinations with Internal Differences of 2 or 3 (i.e., Extreme/Moderate, Extreme/ Strategic and Immoderate/Strategic).

Similar to Distribution 2, the basis for this ranking of probabilities is that the Parties are aware that the method for determining the SPROR rewards those with smaller Internal Differences and penalizes those with larger Internal Differences. If the Parties are even more concerned about minimizing the Internal Difference than they are in Distribution 2, then they will rank those Strategy Combinations with Internal Differences of 0 ahead of those Strategy Combinations with Internal Differences of 1. As in Distribution 2, Strategy Combinations with Internal Differences of 2 or 3 will be assigned the lowest probabilities.

Consistent with the rules for assigning probabilities and the ranking described here, we have assigned a 13% weight to each of the four Intervenor Strategy Combinations with an Internal Difference of 0. We assigned a 9% weight to each of the three Strategy Combinations with an Internal Difference of 1 and a 7% weight to each of the three Strategy Combinations with an Internal Difference of 2 or 3. The sum of these probabilities is 100% (= (13% x 4) + (9% x 3) + (7% x 3)).

Distribution 4

Probability Distribution 4 attaches a 0% probability to those Strategy Combinations which include a Strategic recommendation. The remaining Strategy Combinations without Extreme recommendations (i.e., Immoderate/Immoderate, Immoderate/ Moderate and Moderate/Moderate) are assigned probabilities of 20% each; and the Strategy Combinations with Extreme recommendations (i.e. Extreme/Extreme, Extreme/Immoderate and Extreme/Moderate) are assigned probabilities of 13-1/3% each. The sum of these probabilities is 100% (= (20% x 3) + (13-1/3% x 3)).

Distribution 5

Probability Distribution 5 is a "trapezoidal" Distribution in which the probabilities rise to a plateau as the recommendations move away from Extreme and then decline again as they move towards Strategic. The Extreme/Extreme and Strategic/Strategic Strategy Combinations are associated with 6% probabilities; and the probabilities rise by 2% per Strategy Combination until they plateau at 14% for Immoderate/Immoderate and

Immoderate/Moderate. The sum of the probabilities is 100% (= 6% + 8% + 10% + 12% + 14% + 14% + 12% + 10% + 8% + 6%).

Table 4.1 provides a synopsis of the probabilities attached to each Strategy Combination under each of the five Distributions.

Table 4.1

STRATEGY COMBINATION PROBABILITIES BY DISTRIBUTION

	Probability Distributions				
Strategy Combinations	1	2	3	4	5
Extreme/Extreme	10.0%	10.0%	13.0%	13.3%	6.0%
Extreme/Immoderate	10.0	15.0	9.0	13.3	8.0
Extreme/Moderate	10.0	5.0	7.0	13.3	10.0
Extreme/Strategic	10.0	5.0	7.0	0.0	12.0
Immoderate/Immoderate	10.0	10.0	13.0	20.0	14.0
Immoderate/Moderate	10.0	15.0	9.0	20.0	14.0
Immoderate/Strategic	10.0	5.0	7.0	0.0	12.0
Moderate/Moderate	10.0	10.0	13.0	20.0	10.0
Moderate/Strategic	10.0	15.0	9.0	0.0	8.0
Strategic/Strategic	10.0	10.0	13.0	0.0	6.0
Total Probabilities	100.0%	100.0%	100.0%	100.0%	100.0%

4.3 Analyzing the Expected SPROR Adjustments

Table 4.2 shows the Model 41 SPROR Adjustments for each of the 100 possible Outcomes (= the 70 Outcomes *plus* the 30 Exclusion 2 Outcomes shown on the Test Grid in Part 2.3). Each entry in Table 4.2 is a simple average of the SPROR Adjustments across the three Environments. The right-most column, Five Distribution Average, contains the expected SPROR Adjustments for each Utility Strategy Combination averaged across the five probability Distributions.

Table 4.2

OPTIMAL STRATEGY ANALYSIS FOR UTILITIES ASSUMING MODEL 41 TO DETERMINE SPROR

SPROR Adjustments Averaged Across Three Environments and Five Intervenor Response Distributions

	Intervenors										Five Distribution Average
	Ext/Ext	Ext/Immod	Ext/Mod	Ext/Strat	Immod/Immod	Immod/Mod	Immod/Strat	Mod/Mod	Mod/Strat	Strat/Strat	
Utilities											
Ext/Ext	0.00%	(0.42%)	(0.34%)	(0.12%)	0.00%	(0.37%)	(0.15%)	0.00%	(0.19%)	0.00%	(0.16%)
Ext/Immod	0.42%	0.00%	(0.16%)	0.52%	0.37%	0.00%	0.46%	0.19%	0.00%	(0.56%)	0.13%
Ext/Mod	0.34%	0.16%	0.00%	0.42%	0.15%	(0.46%)	0.00%	(0.62%)	(0.39%)	(0.52%)	(0.11%)
Ext/Strat	0.12%	(0.52%)	(0.42%)	0.00%	(0.68%)	(0.44%)	(0.34%)	(0.58%)	(0.36%)	(0.48%)	(0.40%)
Immod/Immod	0.00%	(0.37%)	(0.15%)	0.68%	0.00%	(0.19%)	0.62%	0.00%	0.56%	0.00%	0.05%
Immod/Mod	0.37%	0.00%	0.46%	0.44%	0.19%	0.00%	0.39%	(0.56%)	0.00%	(0.46%)	0.06%
Immod/Strat	0.15%	(0.46%)	0.00%	0.34%	(0.62%)	(0.39%)	0.00%	(0.52%)	(0.31%)	(0.42%)	(0.27%)
Mod/Mod	0.00%	(0.19%)	0.62%	0.58%	0.00%	0.56%	0.52%	0.00%	0.46%	0.00%	0.23%
Mod/Strat	0.19%	0.00%	0.39%	0.36%	(0.56%)	0.00%	0.31%	(0.46%)	0.00%	(0.36%)	(0.06%)
Strat/Strat	0.00%	0.56%	0.52%	0.48%	0.00%	0.46%	0.42%	0.00%	0.36%	0.00%	0.26%

The data in Table 4.2 show that the optimal Utility Strategy Combination is Strategic/Strategic, which produces an expected, probability-weighted SPROR Adjustment of *plus* 26 basis points. There is no certainty respecting the 26 basis points, because the actual recommendations of the Intervenors are unknown; however, the optimal Utility strategy given this lack of certainty is Strategic/Strategic, followed by Moderate/Moderate with an expected, probability-weighted SPROR Adjustment of *plus* 23 basis points.

4.4 Optimal Strategy and Assumed Probability Distribution

In Part 4.3, we concluded that the optimal Utility Strategy Combination (i.e., the Strategy Combination with the largest positive expected SPROR Adjustments) was Strategic/ Strategic. This analysis was based on average SPROR Adjustment values across three Environments and five alternative Intervenor response probability Distributions.

In Part 4.4, we examine the sensitivity of the optimal strategy to the choice of assumed probability Distribution and Environment. Table 4.3 shows the expected SPROR Adjustments arising from choices of Strategic/Strategic or Moderate/Moderate assuming each of the five probability Distributions averaged across the three Environments. In each case, these values are the *highest* positive SPROR Adjustments for each set of Distributions, confirming that the optimal Utility Strategy Combinations are Strategic/Strategic and Moderate/Moderate irrespective of the assumed Distribution of Intervenor Strategies.

Table 4.3

EXPECTED SPROR ADJUSTMENTS FOR UTILITY STRATEGIES REPORTED BY INTERVENOR RESPONSE PROBABILITY DISTRIBUTION

	Strategic Strategic	**Moderate Moderate**
Distribution 1 · Equal Probability Distribution	0.28%	0.26%
Distribution 2 · ID = 1 Highest Probability ID = 2, 3 Lowest Probability	0.28	0.21
Distribution 3 · ID = 0 Highest Probability ID = 2, 3 Lowest Probability	0.22	0.20
Distribution 4 · Strategic = 0% Probability Immoderate and Moderate are Highest Probability	0.24	0.17
Distribution 5 · Trapezoidal Distribution	0.30	0.30
Average of Five Distributions	0.26	0.23

The principal conclusions that we draw from Table 4.3 are:

1. Averaging Outcomes from the three Environments, the Utility strategies closest to Forecast and farthest from Extreme with the lowest Internal Differences (i.e., Strategic/Strategic and Moderate/Moderate) are statistically expected to yield the optimal results.

2. The qualitative conclusion that Strategic/Strategic and Moderate/Moderate are the optimal strategies for the Utilities is largely unaffected by the assumed Distribution of Intervenor strategies.

4.5 Optimal Strategy and Environment

In addition to considering whether the assumed Distribution of Intervenor strategies affects the potential impact of changing the assumed Distribution of Intervenor strategies, we also considered the possibility that different optimal strategies might exist in different Environments. The three Environments considered in our studies are:

1. Environment 1 · Reality Gap = 0% · Strategy Gap = 0.25% · Maximum Recommendation Spread = 2.0%

2. Environment 2 · Reality Gap = 1% · Strategy Gap = 0.25% · Maximum Recommendation Spread = 3.0%

3. Environment 3 · Reality Gap = 0% · Strategy Gap = 0.5% · Maximum Recommendation Spread = 4.0%

Our *a priori* expectation is that the benefit of adopting a modest Strategy Combination will increase with the Maximum Recommendation Spread.

Table 4.4

EXPECTED SPROR ADJUSTMENTS REPORTED BY ENVIRONMENT

	Strategic **Strategic**	**Moderate** **Moderate**
Environment 1 · Maximum Recommendation Spread = 2.0%[6]	0.16%	0.20%
Environment 2 · Maximum Recommendation Spread = 3.0%[7]	0.32	0.26
Environment 3 · Maximum Recommendation Spread = 4.0%[8]	0.31	0.22
Average of Three Environments	0.26	0.23

The principal conclusions that we draw from Table 4.4 are:

1. The weakest incentives to adopt a reasonable (i.e., non-Extreme) Strategy exist in Environment 1 where the Reality Gap is 0%, the Strategy Gap is 0.25% and the Maximum Recommendation Spread is 2.0%.

2. The strongest incentives to adopt a reasonable Strategy exist in Environment 2 where the Reality Gap is 1%, the Strategy Gap is 0.25% and the Maximum Recommendation Spread is 3.0%.

3. In Environment 3 where the Reality Gap is 0%, the Strategy Gap is 0.5% and the Maximum Recommendation Spread is 4.0%, the optimal strategy

[6] The optimal strategies in Environment 1 are Extreme/Extreme and Immoderate/Immoderate with expected SPROR Adjustments of 0.28% and 0.24% respectively.
[7] The optimal strategy for Environment 2 is Strategic/Strategic, followed by Moderate/Moderate.
[8] The optimal strategy for Environment 3 is Strategic/Strategic, followed by Extreme/Immoderate (expected SPROR Adjustment of 0.29%), followed by Moderate/Moderate. The choice of an Extreme/Extreme Strategy is expected to be disastrous, resulting in a *negative* SPROR Adjustment of 46 basis points.

> is Strategic/Strategic, indicating a strong incentive to adopt a reasonable Strategy. The only concern is that an Extreme/Immoderate Strategy in this Environment edges out a Moderate/Moderate Strategy for "second place" (i.e., an expected positive SPROR Adjustment of 0.29% for Extreme/ Immoderate v. 0.22% for Moderate/Moderate).

If we were speculating about the actual Environment in which recommendations have been recently developed, we would give little or no weight to Environment 1, because *actual* Recommendation Spreads for at least two decades have exceeded its 2.0% maximum (see Chart 1.2). Moreover, if the Parties reduced their Recommendation Spreads to 2.0% or less, then the Commission's frustration would largely disappear, obviating the necessity for further incentives.

Considerations respecting Environments 2 and 3 are mixed. We believe it likely that there exists a genuine Reality Gap between the Parties. Only Environment 2 provides for a Reality Gap, suggesting that greater weight should be placed on the Environment 2 results. On the other hand, the Maximum Recommendation Spread in Environment 2 is 3.0%; and recent Recommendation Spreads have been in the 3.0-4.0% range. This consideration suggests that greater weight should be placed on the Environment 3 results.[9]

For these reasons, we give equal weight to the results from Environments 2 and 3. The optimal Strategy Combinations in Environment 2 are Strategic/Strategic and Moderate/ Moderate; and the optimal Strategies in Environment 3 are Strategic/Strategic and Extreme/Immoderate, with Moderate/Moderate in third place.

4.6 Conclusions

Chapter 4 examines optimal strategies for the Parties if the Commission adopts Model 41 as the basis for determining the Starting Point Rate of Return (SPROR). Consistent with the studies and conclusions in Chapter 3, the optimal strategies derived in Chapter 4 are, for the most part, strategies which avoid Extreme recommendations and embrace Moderate or Strategic recommendations. This conclusion applies across a range of assumed probability Distributions respecting the potential actions of the other Party and in the circumstances of Environments that are closest to recent experience.

In plain language, if we were advising a Utility faced with a regulatory regime that used Model 41 to establish a Starting Point Rate of Return, we would recommend an initial analysis of changes in general economic circumstances, including inflation and long-term bond yield trends, and the logical application of these changes to the most recently-awarded common equity rate of return to establish a realistic view of the Forecast rate of return. We would then urge the Utility to proffer rate of return evidence with

[9] Even so, if the incentives that we propose here have their intended effect, then Recommendation Differences will hopefully decline below 3.0%.

recommendations no higher than 50 basis points above the Forecast rate of return and preferably with an Internal Difference of 0-25 basis points.

If we were advising an Intervenor faced with the same regime, we would recommend the same care and attention to estimating the Forecast rate of return. We would then urge the Intervenor to proffer rate of return evidence with recommendations no lower than 50 basis points below the Forecast rate of return and preferably with an Internal Difference of 0-25 basis points.

If both Parties accepted our recommendations – based on the research set out here – then the Recommendation Spread would be no more than 2.0% assuming a Reality Gap of no more than 1.0%.

Chapter 5

APPLYING MODEL 41 TO SEVEN COMMISSION DECISIONS

To our knowledge, the universe of litigated, multi-company rate of return decisions by the Commission for major utilities consists of the seven decisions from *Decision U97065* to *Decision 20622-D01-2016*.[1] The incentives inherent in our recommended Model 41 for determining the Starting Point Rate of Return (SPROR) were not present when the Parties proffered their recommendations in these proceedings and when the Commission made its decisions. Nevertheless, we were irresistibly curious about the question: *What results would Model 41 have yielded if it had been applied to the recommendations in each of the seven proceedings?*

Of course, if the Parties had known in advance that the Commission intended to apply Model 41 to determine an SPROR, then the recommendations may have been different – in fact, we hope that the recommendations would have been more moderate. Therefore, our purpose in applying the recommended Model to these historical data is to ascertain whether the results are so materially different from the final decision that they would render the approach unrealistic and completely unacceptable to the Commission or one or both of the Parties.

Table 5.1 shows the calculated SPRORs for each of the seven proceedings, the average of the four recommendations in each proceeding and the Commission's decision.

Table 5.1

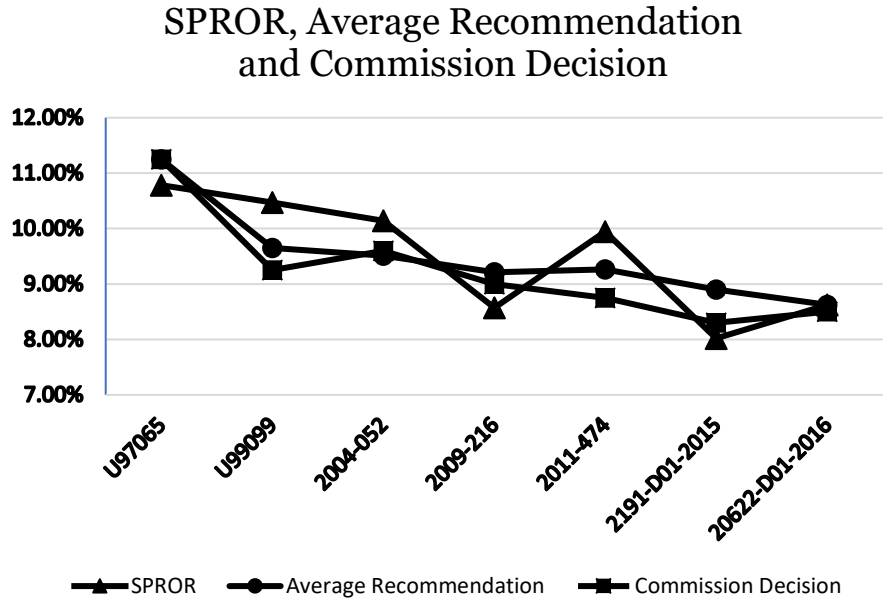
STARTING POINT RATES OF RETURN, AVERAGE RECOMMENDATIONS AND AWARDED RATES OF RETURN

Decision	SPROR	Average Recommendation	Awarded Rate of Return
U97065	10.78%	11.25%	11.25%
U99099	10.47	9.65	9.25
2004-052	10.14	9.51	9.60
2009-216	8.57	9.21	9.00
2011-474	9.94	9.26	8.75
2191-D01-2015	8.02	8.90	8.30
20622-D01-2016	8.63	8.63	8.50
Averages	9.51%	9.49%	9.24%

[11] The decisions are: *Decision U97065*, *Decision U99099*, *Decision 2004-052*, *Decision 2009-216*, *Decision 2011-474*, *Decision 2191-D01-2015* and *Decision 20622-D01-2016*.

The data from Table 5.1 are presented on Chart 5.1

Chart 5.1



We draw three conclusions from the data in Table 5.1 and Chart 5.1. First, the average SPROR and the average of each proceeding's average recommendation are virtually the same, although differences do exist in most of the years. A reasonable inference is that if the Commission had given significant weight to the SPROR in each of the proceedings, then it may have reached a materially different result on a decision-by-decision basis but is unlikely to have reached a materially different average result for the seven decisions considered as a whole.

Second, the absolute values of the decision-by-decision differences between the SPROR and the Average Recommendation range from zero to 88 basis points. In Part 2.6, we indicated that SPROR Adjustments whose absolute values exceeded 125 basis points or 100 basis points for "non-Extreme" Strategies only were likely to be unacceptable to the Commission or the Parties; and, as a result, we rejected Models whose maximum SPROR Adjustments exceeded these limits. The range of SPROR Adjustments from Model 41 applied to the seven decisions is well within these ranges.

Third, the data in Table 5.1 raise no concerns about the use of Model 41 to determine the SPROR. If Parties are aware that the Commission intends to use the recommended Model, then the results are likely to reflect a greater degree of moderation than the results that we derive from this "back-casting" exercise. If the Parties were to be more or less equally influenced by the pressure towards moderation, then it is likely that the Commission's decision would not be materially different; however, the Recommendation Spread would likely be smaller. If the Parties were not more or less equally influenced by the pressure towards moderation, then the Party who exhibited the greater degree of

moderation (i.e., the less extreme Party) would be likely to enjoy an advantage over the more extreme Party.

Chapter 6

CONCLUSIONS, RECOMMENDATIONS AND FURTHER STUDY

Recommended common equity rates of return for public utilities are the product of expert studies and evidence which, of necessity, require professional judgment. Unfortunately, the ranges of recommendations have widened to 3-4% in recent years, thus limiting their usefulness to the Commission. This phenomenon may have arisen because the Commission's rate of return decisions have tended to be close to the midpoint of the competing recommendations, thereby reinforcing the wisdom of parties in taking ever more extreme positions in an effort to "pull the average up" or "push the average down." This frustrating situation is discussed in Chapter 1.

In Chapter 2, we designed tests for evaluating alternative Models to encourage Utilities and Intervenors to moderate their more extreme positions and "move to the middle." The tests are applied to 66 Models in Chapter 3, where we conclude that the most effective of the 66 Models is one in which the Starting Point Rate of Return (SPROR)[1] is calculated by reference to a weighted average of the recommendations which gives greatest weight to recommendations closest to the average and least weight to recommendations furthest from the average.[2]

In Chapter 4, we investigated the optimal strategy that Parties should adopt given the expectation that the Commission intends to give weight to a SPROR determined on the basis of the recommended Model. The analysis in Chapter 4 confirms that Parties have a clear incentive under the recommended Model to "move to the middle" and moderate their rate of return recommendations.

We recommend that the Commission announce that it intends to set aside any notion of "splitting the difference" in rates of return in favour of a formula for determining a Starting Point Rate of Return that will reward Parties for moderating their recommendations. The Commission's final decision respecting rate of return may or may not be coincident with a mechanistically-determined SPROR, because there is no certainty that the SPROR will satisfy the legal requirements for a fair return. Nevertheless, the Parties can expect that the SPROR is likely to receive considerable weight in the determination of the fair return. And if the Parties hold this belief, then they will logically respond by seeking a middle ground and avoiding extreme positions.

[1] The Starting Point Rate of Return is the rate of return broadly indicated by the recommendations on the record – i.e., the "starting point" for the Commission's consideration of the fair return question.
[2] A more precise description of the recommended Model, set out in Part 3.5, includes an adjustment for Internal Differences in the calculation of the differences between each recommendation and the average recommendation.

We look forward to the possibility of expanding our current research to possibly incorporate:

1. the mathematical derivation of a generalized solution for Models which cause the Parties to "move to the middle"

2. generalization of the current work to include Outcomes with varying numbers of Utility and Intervenor recommendations – i.e., not simply two recommendations from each of the Parties

3. a mathematical proof that the optimal Fulcrum is 2.25%

4. the possible testing of other Environments

5. the possible testing of other Intervenor response probability distributions

John P. Evans
Antigonish, Nova Scotia

Robert E. Evans
Calgary, Alberta

John P. Evans is a fourth-year Honours Economics student at St. Francis Xavier University. Mr. Evans is President of the St. FX Economics Society and a Founder and Editor-in-Chief of *X Markets Monthly*, a student investment publication jointly sponsored by the Department of Economics and the Schwartz Business School. Mr. Evans has recently been engaged in independent economic research on the impact of abandoned coal mines on methane production in Nova Scotia. Mr. Evans' interests include the application of game theory to real-world political and economic problems.

Robert E. Evans is President of Economic Research Associates Limited, a consulting firm specializing in problems of the regulated electric, gas, pipeline and transportation industries. Dr. Evans has studied regulatory problems for over 40 years and has appeared as a witness on rate of return and other matters of regulatory finance in proceedings before regulators or courts in Canada, New Zealand and the United States.

GLOSSARY

Close Outcome · A Close Outcome is an Outcome in which the rewards or penalties are too weak to be considered meaningful. We specifically define a Close Outcome as an Outcome with an SPROR Adjustment whose absolute value is less than 10 basis points.

Commission · The Alberta Utilities Commission

Distribution · Distribution refers to the distribution of probabilities respecting each of the possible Strategy Combinations that might be adopted by Utilities or Intervenors. Five Distributions are considered and described in Part 4.2.

Environment · An Environment is a combination of a Reality Gap and a Strategy Gap which defines the rates of return associated with each Strategy. Three Environments are considered and described in Part 2.1.

Extreme · Extreme is the Strategy which is furthest from Forecast within each Environment for each Party – i.e., the *highest* Utility rate of return recommendation or the *lowest* Intervenor recommendation in each Environment.

Fairness · Fairness means that a symmetric end result will occur irrespective of whether the Strategy Combinations are those embraced by Utilities or Intervenors.

Fixed Sum Method · A method for assigning weights in a WAFA Model where the sum of the weights attached to each recommendation is fixed at either 9 or 11, depending on whether the minimum difference recommendation receives a weight of 4x or 6x. With the Fixed Sum Method, the weights of, say, 4, 2 and 1 (or 6, 2 and 1) must sometimes be adjusted if more than one recommendation is either closest to or further from the average. The Fixed Sum Method is described in Part 2.4.

Fixed Value Method · A method for assigning weights in WAFA Model where the weights attached to each recommendation never vary, irrespective of the fact that it is possible for more than one recommendation to be "tied" for closest or furthest from the central value. For example, if two recommendations are "tied" for closest to the central value and if the WAFA Model under consideration weights the closest recommendation 4x, then there would be two recommendations that would receive a weight of 4x. In contrast, the Fixed Sum Method would assign weights of 3x to these recommendations in order to maintain a fixed sum of 9.

Forecast · Forecast is the Strategy that reflects each Party's best private estimate of the common equity rate of return that will actually be awarded by the Commission.

Fulcrum · If the Internal Recommendation Spread is *below* the Fulcrum, then the Internal Difference Adjustment is *added* to the absolute values of the differences between each recommendation and the central value for all recommendations. If the Internal Recommendation Spread is *above* the Fulcrum, then the Internal Difference Adjustment is *subtracted* from the absolute values of the differences between each recommendation the central value for all recommendations. If the Internal Recommendation Spread equals the Fulcrum, then no adjustment is made. In our work, we use Fulcrums of 2.0%, 2.25% and 2.5%.

Game · In the present context, a Game is a set of rules that define how the Starting Point Rate of Return (SPROR) will be determined. The words "Game" and "Model" are synonymous.

GCOC · Generic Cost of Capital · A proceeding in which a fair common equity rate of return is determined for an illustrative utility of average risk.

ICE Models · A variation on the WAFA Model where the weights assigned to each of the recommendations and Internally-Calculated Exponential (ICE) weights. The mechanics of the weight calculations are described in Part 2.4. Models 49 – 66 are ICE Models.

Immoderate · Immoderate is the Strategy closest to Extreme. The Immoderate rate of return within each Environment is the Extreme rate of return adjusted by the Strategy Gap. For example, if the Utility Extreme rate of return is 9.5% and if the Strategy Gap is 0.25%, then the Utility Immoderate rate of return is 9.25%. Or if the Intervenor Extreme rate of return is 7.5% and if the Strategy Gap is 0.25%, then the Intervenor Immoderate rate of return is 7.75%.

Initial Models · 24 Models that reflect our initial thinking about how to encourage Parties to "move to the middle." The 24 Models use six Structures, two Simple Average Limits and either include or exclude an Internal Difference Adjustment.

Internal Difference · The difference between the rate of return recommendations of two Utility experts or between the rate of return recommendations of two Intervenor experts. For example, if the Utility experts recommend rates of return of 10.0% and 9.5%, then the Internal Difference is 0.5% (= 10.0% *less* 9.5%). Internal Differences may also be expressed as the number of Strategies separating the two recommendations of either Utilities or Intervenors. To illustrate, if the Intervenors proffer rates of return which are Immoderate and Strategic, then the Internal Difference can be expressed as a 2 (i.e., Immoderate to Moderate and Moderate to Strategic). When an Outcome is said to have Internal Differences of 2/1 this means that the Utilities have an Internal Difference of 2 (e.g., Extreme/Moderate); and the Intervenors have an Internal Difference of 1 (e.g., Moderate/Strategic).

Internal Difference Adjustment · An adjustment to the absolute values of the differences between each recommendation and the central value (either average or midpoint) for all recommendations. In those Initial Models where an Internal Difference Adjustment is

made, the Internal Difference is always *added* to the absolute values of the differences. In WAFA Models, the Internal Difference Adjustments are added or subtracted from the absolute values of the differences, depending on the qualitative relationship between the Fulcrum and the Internal Recommendation Spread (see Fulcrum).

Internal Recommendation Spread · The difference between the *lowest* Utility rate of return recommendation and the *highest* Intervenor rate of return recommendation.

Intervenors · Entities whose interests are generally furthered by minimizing the rate of return awarded by the regulator. Customer groups and consumer advocates are examples of prominent Intervenors.

Lose · Lose for the Utilities means an SPROR *lower* than the average of the four recommendations (i.e., a *negative* SPROR Adjustment). Lose for the Intervenors means an SPROR *higher* than the average of the four recommendations (i.e., a *positive* SPROR Adjustment).

Maximum Recommendation Spread · The Maximum Recommendation Spread for a particular Environment is the percentage difference between the rate of return that is Extreme for Utilities and the rate of return that is Extreme for Intervenors. The Maximum Recommendation Spreads for the three Environments that we considered are 2.0%, 3.0% and 4.0%.

Model · See "Game"

Moderate · Moderate is the Strategy halfway between Extreme and Forecast. If the Utility Extreme rate of return is 9.5% and if the Utility Forecast rate of return is 8.5%, then the Utility Moderate rate of return is 9.0%. Or if the Intervenor Extreme rate of return is 7.5% and if the Intervenor Forecast rate of return is 8.5%, then the Intervenor Moderate rate of return is 8.0%.

Negative Incentive · A Negative Incentive arises when extremity is meaningfully rewarded and moderation is penalized. Negative Incentive is one of the four qualitative results that can result from the tests of the 66 Models.

NIPTO · Negative Incentive as a Percent of Total Outcomes · The ratio of the number of Negative Incentive results to the 210 total Outcomes for each of the 66 Models.

Outcome · A set of four Strategies, two of which are derived from recommendations by Utility experts and two of which are derived from recommendations by Intervenor experts. An Outcome may also be thought of as the four Strategies implicit in a Utility Strategy Combination and an Intervenor Strategy Combination (see definition of Strategy Combination).

Parties · Utilities and Intervenors

PIPTO · Positive Incentive as a Percent of Total Outcomes · The ratio of the number of Positive Incentive results to the 210 total Outcomes for each of the 66 Models.

POOA · Percent of Outcomes Other Than Average · The ratio of the sum of Positive Outcomes and Negative Outcomes to the 210 total Outcomes for each of the 66 Models.

Positive Incentive · A Positive Incentive arises when moderation is meaningfully rewarded and extremity is penalized. Positive Incentive is one of the four qualitative results that can result from the tests of the 66 Models.

Reality Gap · Any difference which may exist between the private perceptions of Utilities and Intervenors respecting the common equity rate of return which the Commission will ultimately award (i.e., the rate of return associated with each Party's Forecast Strategy).

Recommendation Spread · The difference between the *highest* Utility rate of return recommendation and the *lowest* Intervenor rate of return recommendation – i.e., the range which encompasses *all* recommendations.

Simple Average Limit · The percentage which triggers the Simple Average Rule (see Simple Average Rule).

Simple Average Outcome · An Outcome in which the Simple Average Rule applies (see Simple Average Rule)

Simple Average Rule · The Simple Average Rule states that if the Internal Recommendation Spread is less than the Simple Average Limit, then the SPROR is calculated as the simple average of the *lowest* Utility recommendation and the *highest* Intervenor recommendation.

SPROR · Starting Point Rate of Return · The SPROR is the rate of return broadly indicated by the recommendations on the record – i.e., the "starting point" for the Commission's consideration of the fair return question.

SPROR Adjustment · The difference between the Starting Point Rate of Return and the simple average of the four rate of return recommendations associated with each Outcome.

Starting Point Rate of Return · The Starting Point Rate of Return or SPROR is the rate of return broadly indicated by the recommendations on the record – i.e., the "starting point" for the Commission's consideration of the fair return question.

Strategic · Strategic is the Strategy closest to Forecast. The Strategic rate of return within each Environment is the Forecast rate of return adjusted by the Strategy Gap. For example, if the Utility Forecast rate of return is 9.0% and if the Strategy Gap is 0.25%, then the Utility Strategic rate of return is 9.25%. Or if the Intervenor Forecast rate of return is 8.0% and if the Strategy Gap is 0.25%, then the Intervenor Strategic rate of return is 7.75%.

Strategy · A name given to recommended rates of return that are set at specified distances from the rate of return that Parties privately forecast that the Commission is most likely to award. The Strategy names are Forecast, Strategic, Moderate, Immoderate and Extreme.

Strategy Combination · Two Strategies adopted by the same Party. For example, Extreme/Extreme or Moderate/Strategic are Strategy Combinations that could be adopted by either the Utilities or the Intervenors.

Strategy Gap · The difference between common equity rates of return as Parties move from one Strategy to the next closest Strategy. To illustrate, if the Utility's Immoderate Strategy is associated with a 9.25% rate of return and if the Utility's Moderate Strategy is associated with a 9.0% rate of return, then the Strategy Gap is 0.25% (= 9.25% *less* 9.0%).

Structure · There are six Structures associated with the 24 Initial Models. Each Structure uses a different method for determining the SPROR. To illustrate, the simplest Structure is to take that recommendation closest to the average of the four recommendations and adopt that closest recommendation as the SPROR. An alternative Structure is to compute the standard deviation of the four recommendations, exclude any recommendations more than one standard deviation removed from the mean and then recalculate the average as the SPROR.

Summary of Results · A table for each of the 66 Models described in Part 3.1 which displays the test results for that Model. Table 3.1 is the Summary of Results for Model 41.

Utilities · Companies whose earnings are subject to rate of return regulation. Under most circumstances, Utilities have an incentive to maximize the rate of return awarded by the regulator.

WAFA Models · Weighted Average Fulcrum Adjustment Models · Models 25 – 48 are WAFA Models in which a Fulcrum determines whether an Internal Difference Adjustment will be *added* or *subtracted* from the absolute values of the differences between each recommendation and the central value for all recommendations to determine the weights for each recommendation in the SPROR calculation.

Win · Win for the Utilities means an SPROR *higher* than the average of the four recommendations (i.e., a *positive* SPROR Adjustment). Win for the Intervenors means an SPROR *lower* than the average of the four recommendations (i.e., a *negative* SPROR Adjustment).

APPENDIX

John P. (Jack) Evans

2032 Palisprior Road, SW · Calgary, Alberta T2V 5J7
587-718-2032 | jackevans@shaw.ca

EDUCATION

St. Francis Xavier University **Antigonish, Nova Scotia**
Bachelor of Arts in Economics (Honours) Expected June 2018

- **Relevant Coursework:** Corporate Finance (2 courses), Investments, Financial Accounting, Econometrics, Monetary Policy, Micro and Macro Economic Theory (8 courses), Calculus (3 Courses), Game Theory and Real Analysis (1 course), Matrix and Linear Algebra (2 courses), Financial Mathematics (1 course), and Financial Modelling (1 course).
- **Professional Designations:** Currently enrolled in the Canadian Securities Course.
- **Teaching Assistant:** Introductory Macroeconomics
- **Grades:** 3.8 GPA (3.9 GPA in Year 3), Honours Designation, Dean's List 2014-2015, 2015-2016, 2016-2017

Rundle College High School **Calgary, Alberta**
Honours Graduate June 2014

WORK & LEADERSHIP EXPERIENCE

St. Francis Xavier University Economics Department **Antigonish, Nova Scotia**
Research Assistant May 2017 – Present

- Selected Work Experience
  - Natural Resources Canada (NRCan) methane gas seepage project (GASP).
  - Examining methane emissions from retired Nova Scotia coal mines and New Brunswick oil fields.
  - Tasked with building a comprehensive cost-benefit analysis model to determine mitigation economics.
  - Work includes: literature review, excel modelling, sensitivity analysis, macroeconomic analysis, marginal abatement cost (MAC) curve construction, McKinsey curve construction, emissions projections, discount rate determination, and policy recommendations.
  - Provide marginal cost estimates of emissions for research and development of mitigation methods in project's Phase II.

St. Francis Xavier University Economics Society **Antigonish, Nova Scotia**
President September 2016 – May 2017

- Selected Work Experience
  - Tasked with running weekly meetings and representing the Society at faculty and student functions.
  - Recruited record numbers of members this year.
  - Organized speakers, student tutoring, and working to institute an economics alumni network.

Wild Wave Watercraft and Boat Rentals **Bigfork, Montana USA**
Manager, Bigfork Operations Summers June 2010 – August 2016

- Managed employees, watercraft fleet and gas pump for Wild Wave resulting in increased customer loyalty and increased company revenues.
- Selected Work Experience
  - Manage staff, prepare work schedules, and assign specific duties.
  - Determine staffing requirements, interview, hire, and train new employees.
  - Rent and maintain a fleet of 15 boats and 8 Seadoos.
  - Building and implementation of organizational computer software to streamline business operations and increase communication.
  - Promoted to manager in the summer of 2014.

St. Francis Xavier University Student's Union **Antigonish, Nova Scotia**
Vice-President, Riley Hall (Student Residence) April 2015 –April 2016

- Riley Hall finished the year with a record number of house points, the most organized house council and a Book of Honour to preserve residence history.
- Selected Work Experience
    - Organize and facilitate house council meetings using Robert's Rules of Order.
    - Responsible for creating and distributing weekly meeting agendas.
    - Responsible for taking and distributing meeting minutes.
    - Responsible for all budget and accounting matters for Riley Hall.

SKILLS, ACTIVITIES & INTERESTS

Technical Skills: Corporate certified in advanced Microsoft Excel and intermediate Microsoft Access by CTS training (Chicago). Proficient in Stata regression software and programing in MATLAB.
Illustrative Activities: St. FX Volleyball Society, Shineorama (CF fundraising), and actor for Theatre Antigonish.
Professional Interests: Interested in global financial markets, real-life securities analysis, economic development, and intent on pursuing a doctorate degree in financial economics.
Personal Interests: Golf, basketball, volleyball, acting, cooking, music (violin and trumpet), travel, and writing.
Books Recently Read: *Dual Momentum Investing*, *How to Win Friends and Influence People*, and *The Bottom Billion*
Currently Reading: *A Random Walk Down Wall Street*
Additional Information: North American First Nation/Indian heritage- Enrolled Member of the Cheyenne River Sioux Nation.

PROFESSIONAL QUALIFICATIONS OF ROBERT E. EVANS

Academic Training

Ph.D. in Economics from University of Wisconsin-Madison (1980)
Major Field: Industrial Organization
Minor Fields: Finance and Econometrics
Third Field: Money and Banking

M.S. in Economics from University of Wisconsin-Madison (1976)

B.S. in Economics, *magna cum laude*, Washington and Lee University (1974)

Professional and Academic Organizations

Omicron Delta Epsilon, Honorary Economics Society
Phi Beta Kappa, Honorary Scholastic Society
Phi Eta Sigma, Honorary Scholastic Society

Professional Experience

September 1976 - July 1977 · Employed by the Rates Division of the Public Service Commission of Wisconsin · Madison, Wisconsin

October 1977 - October 1980 · Consultant, Foster Associates, Inc. · Washington, D. C.

February 1981 - November 1982 · Vice-President and Consultant, Pitfield Mackay Ross Ltd. · Toronto

November 1982 - 1998 · Consultant, Economic Research Associates Limited · Toronto

1998 – present · President, Economic Research Associates Limited · Calgary

Current Contact Details

Economic Research Associates Limited
Box 72008 · D167, 1600 90th Avenue, SW
Calgary, Alberta T2V 5H9
403.708.2623 · ree@shaw.ca

Previous Evidence

Dr. Evans has given evidence on one or more occasions before the following Canadian tribunals and in respect of the following companies:

Public Utilities Board, Alberta
Alberta Energy and Utilities Board · Alberta Utilities Commission

AltaLink, L. P.
Aquila Networks Canada (Alberta) Ltd.
Edmonton Power
Edmonton Power Generation Inc.
Edmonton Power Transmission Inc./EPCOR Transmission Inc.
EPCOR Distribution Inc.
EPCOR Distribution and Transmission Inc.
EPCOR Energy Alberta Inc. / EPCOR Energy Alberta, GP Inc.
Grande Prairie Transmission Company
NOVA Corporation of Alberta
NOVA Gas Transmission Ltd.
TransAlta Utilities Corporation

British Columbia Utilities Commission

Columbia Natural Gas Limited
Fort Nelson Gas Ltd.
Inland Natural Gas Co. Ltd.
West Kootenay Power

Canadian Radio-television and Telecommunications Commission

British Columbia Telephone Company
Telesat Canada

Canadian Transport Commission

Cost of Capital for Various Purposes Under the *Railway Act* and the *Western Grain Transportation Act*

Public Utilities Board of Manitoba

Greater Winnipeg Gas
Steelgas Utilities of Manitoba

National Energy Board

Alberta Natural Gas Company Ltd
Foothills Pipe Lines (Yukon) Ltd.
Interprovincial Pipe Line Limited
Trans Mountain Pipe Line Company Ltd.
Trans-Northern Pipelines Inc.

New Brunswick Board of Commissioners of Public Utilities

The New Brunswick Telephone Company Limited

Northwest Territories Public Utilities Board

Stittco Utilities (NWT) Ltd.

Province of Nova Scotia Board of Commissioners of Public Utilities

Maritime Telegraph and Telephone Company Limited
Nova Scotia Power Incorporated

Ontario Energy Board

Northern and Central Gas Corporation

Prince Edward Island Public Utilities Commission

The Island Telephone Company Limited
Maritime Electric Company Limited

Other Interests

Nia Technologies Ltd. · Director and Secretary · 2015 – present

Christian Blind Mission International · Director, 2006 – 2015 · Secretary, 2010 - 2015

Rundle College Society · Director, 2001 - present · Vice-Chairman, 2003-2004 · Chairman, 2004 – 2016

Suzuki Talent Education Society · Director, 2000 – 2006 · Treasurer, 2002 - 2006