INDEX

TESTIMONY of

JUERGEN M. BERMEJO and KATHERINE L. BEALE

Witnesses for Bonneville Power Administration

SUBJECT: VARIABLE COST PRICING METHODOLOGY

Page

Section 1: Introduction And Purpose Of Testimony 1

Section 2: Description of the Proposed Variable Cost Methodology.....2

Section 3: Specific Variable Costs Considered By The GARD Model 8

Section 4: Apportioning the Variable Cost of Reserves.....24

Section 5: Impacts of Changing the Persistence Scheduling Assumptions on the Variable Cost Allocation 28

This page intentionally left blank.

TESTIMONY of

JUERGEN M. BERMEJO and KATHERINE L. BEALE

Witnesses for Bonneville Power Administration

SUBJECT: VARIABLE COST PRICING METHODOLOGY

SECTION 1: INTRODUCTION AND PURPOSE OF TESTIMONY

Q. Please state your names and qualifications.

A. My name is Juergen M. Bermejo, and my qualifications are contained in WP-10-Q-BPA-05.

A. My name is Katherine L. Beale, and my qualifications are contained in WP-10-Q-BPA-04.

Q. What is the purpose of your testimony?

A. The purpose of this testimony is to sponsor section 4 of the Generation Inputs Study (Study), WP-10-E-BPA-08, which describes the variable cost incurred by BPA when it is providing capacity for Regulating Reserve, following reserve, imbalance reserve, and Operating Reserve, and to describe the allocation of these variable costs to Transmission Services (TS) for these services.

Q. How is your testimony organized?

A. The testimony is organized into five sections. Section 1 is the introduction. Section 2 describes the proposed variable cost methodology. Section 3 describes the specific variable costs considered by the Generation and Reserves Dispatch (GARD) model. Section 4 describes the apportioning of the variable cost of reserves to reserve type. Section 5 describes the effects of potential changes in the Generation Reserve Forecast on the variable cost allocation forecast.

SECTION 2: DESCRIPTION OF THE PROPOSED VARIABLE COST METHODOLOGY

Q. Why are you proposing to include a variable cost component in its generation inputs cost allocation?

A. BPA operates an interconnected system of dams and reservoirs, and providing reserves affects system output in MWs, timing of energy generated, and revenues received. Losses of efficiency and value occur as the system is set up to allow reserves to be deployed, and additional losses occur as the reserves are actually deployed. Including this variable cost component appropriately allocates these losses to the parties who benefit from reserve services.

Q. Generally, how is the variable cost of reserves calculated?

A. As discussed in the Study, WP-10-E-BPA-08, section 4, the variable cost of reserves is calculated in two general steps. The first step calculates the cost of making Federal Columbia River Power System (FCRPS) capability ready and available should the need to deploy reserves arise. These costs are referred to throughout this document as stand ready costs. The second step calculates the cost of actually deploying reserves as the need arises. These costs are referred to throughout this testimony as deployment costs. The tool used to calculate these costs is the GARD model. The GARD model is described in detail in the Study, WP-10-E-BPA-08, section 4.

Q. What is the GARD model, and what does it do?

A. The GARD model is an Excel spreadsheet-based model developed by BPA. The GARD model calculates and values the energy shift costs, efficiency loss, and base cycling loss associated with standing ready with the capability to increase or decrease (*inc* or *dec*) generation and the response loss, incremental cycling loss, incremental spill, and incremental efficiency loss associated with deploying reserves.

Q. Please describe the inputs for the GARD model.

A. The GARD model uses inputs from the HYDSIM model, a hydraulic model of coordinated river operations that calculates monthly average generation values by project, and actual system data. The primary inputs into the GARD model are tables of project-specific generation values calculated by HYDSIM.

Q. What are the tables of project-specific generation values used for?

A. These generation tables are used by GARD to determine the generation request and project dispatch. The generation request is the amount of Heavy Load Hours (HLH) or Light Load Hours (LLH) generation that a specific project is being asked to produce. The project's dispatch is the number and/or combination of online units required to meet the generation request and reserve obligation.

Q. How do you determine a project's specific HLH and LLH generation request?

A. HYDSIM output tables are input into a pre-processing spreadsheet to calculate each project's HLH and LLH generation request prior to considering reserve needs. Determining a project's specific HLH and LLH generation request begins with monthly energy amounts for each of the 70 historical water years from HYDSIM. Monthly energy amounts are taken for Grand Coulee (GCL), Chief Joseph (CHJ), John Day (JDA), and The Dalles (TDA). These four projects are generally referred to as controller projects in this testimony. Additionally, the pre-processing spreadsheet calculates amounts of pre-existing *dec* capability for each project by month and historical water year based on the calculated LLH generation and the project's minimum flow. The purpose of pre-existing *dec* capability input is to avoid unnecessarily moving energy out of HLH and into LLH when providing *dec* capability.

Q. Please describe the nature of the generation request in more detail and explain how it is applied to the GARD model.

A. The generation request is a set of generation values, for each month of each water year for HLH and LLH, for each controller project. Given these generation values, the GARD model finds the efficiency-maximizing unit dispatch. This process mimics the basepoint setting process, where the hydro duty scheduler submits requested generation amounts to each project and the project dispatches its units in the most efficient manner possible in order to meet the requested generation.

Q. If the Big 10 projects are all capable of AGC response, why were costs calculated based only on these four controller projects?

A. Although all of the Big 10 projects are capable of being, and at various times of the year are, armed for Automatic Generation Control (AGC) response, GCL, CHJ, JDA, and TDA are the only projects analyzed because these four controller projects are most often armed by the hydro duty scheduler for AGC response. The projects used in the GARD model are representative of the entire Big 10, because the variable costs calculated by the GARD model are based on recovering the total amount of capacity reserves needed for generation inputs provided to TS, and this cost calculation is unlikely to vary to any significant degree whether spread over the four controller projects or all 10 projects. Using the four controller projects that are armed most often makes the GARD model manageable, compared to trying to factor in all 10 projects that have varying degrees of use depending on the time of year.

Q. What is the next step in calculating variable costs?

A. Supervisory Control and Data Acquisition (SCADA) data are used to develop relationships between average energy and HLH generation for each of the respective controller projects. This evaluation is constrained by unit availability, requirements to operate within one percent of peak efficient generation, and minimum turbine flow

constraints. The results of the evaluation are functional relationships between average energy production and HLH generation.

Q. What time period does the SCADA data span?

A. Data are from the period 01/01/02 through 12/31/07.

Q. Why is the 2002 through 2007 dataset used in developing the functional relationship between energy and HLH?

A. The 2002 through 2007 period is used to balance the need for a robust data set with the desire for operations that are similar to current practice and bound by similar constraints. Going back farther in time would include periods when operations were significantly different from current operations. Additionally, this period serves well as a base case because it pre-dates the large wind fleet buildup.

Q. The GARD model refers to incremental (inc) reserves and decremental (dec) reserves; please explain what these are.

A. *Inc* reserves are reserves required to maintain load-resource balance when an under-generation situation exists within the BPA Balancing Authority Area (BAA). In an under-generation situation, instantaneous loads are higher than planned and/or the instantaneous wind generation is lower than planned. Under these circumstances, FCRPS generation must automatically *inc* to maintain system balance. The *inc* reserve is the amount that the FCRPS must be capable of instantaneously increasing generation. The quantity of reserve is set to cover 99.75 percent of all under-generation magnitudes estimated by the Generation Reserve Forecast. Study, WP-10-E-BPA-08, section 2.

Conversely, *dec* reserves are reserves required to maintain load-resource balance when an over-generation situation exists within the BPA BAA. In an over-generation situation, instantaneous loads are lower than planned and/or the instantaneous wind generation is higher than planned. Under these circumstances, FCRPS generation must automatically *dec* to maintain system balance. The *dec* reserve is the amount that the

FCRPS must be capable of instantaneously decreasing generation. The quantity of reserve is set to cover 99.75 percent of all over-generation magnitudes estimated by the Generation Reserve Forecast. Study, WP-10-E-BPA-08, section 2.

Q. Please explain how inc reserves and dec reserves were treated in the GARD model.

A. Reserves are input into the GARD model in the following three categories: 1) the total *inc* spinning obligation inclusive of the spinning portion of the Operating Reserve obligation, 2) the spinning portion of the Operating Reserve obligation, and 3) the *dec* obligation. We will explain each in turn.

First, the spinning portion of the total reserve obligation is explicitly input into the GARD model to ensure maintenance of sufficient total spinning capability at each of the controller projects. The spinning portion of the reserve obligation is the sum of 100 percent of the regulation requirement, 50 percent of the load following requirement, and 50 percent of the total Operating Reserve requirement.

Second, the spinning portion of the Operating Reserve obligation is also input standing alone so the GARD model can identify and track the portion of the total spinning obligation attributable to Operating Reserve. In this way, the GARD model maintains at all times a minimum spinning capability equal to the Operating Reserve obligation during the course of within-hour reserves deployment.

Third, the total *dec* obligation is identified so the GARD model knows how much minimum generation capability is required to provide the reserve. Note that by definition of how the reserve is met, *dec* obligations are spinning.

Q. Why does the GARD model need to maintain the spinning portion of the Operating Reserve obligation at all times as compared to the balancing spinning obligation?

A. Operating Reserve is not used for balancing purposes and may be deployed only for qualifying contingencies. Operating Reserve must be maintained within the BAA equal to five percent of the instantaneous wind and hydro generation and seven percent of the

thermal generation. Because the Operating Reserve requirement is an instantaneously calculated requirement, even momentarily using operating reserve capability for balancing purposes is a violation of Western Electricity Coordinating Council (WECC)/North American Electric Reliability Corporation (NERC) disturbance criteria.

Q. How are non-spinning reserves considered?

A. The amount of reserve that may be carried non-spinning is not directly input; rather, it is calculated from the three reserve input categories described in the preceding paragraph and the total *inc* obligation. From the Generation Reserves Forecast, the 0.9975th percentile of the control error signal distribution is equal to the total *inc* balancing reserve obligation (not including Operating Reserve). Study, WP-10-E-BPA-08, section 2. The total *inc* balancing reserve obligation consists of both a minimum spinning requirement and a non-spinning amount. The difference between the total *inc* balancing reserve obligation and the required *inc* spinning obligation equals the maximum amount of reserve that may be carried as non-spinning. Thus, the difference between the total *inc* balancing reserve obligation and the total *inc* spinning obligation less Operating Reserve is the amount of *inc* balancing that may be carried as a non-spinning reserve.

Q. What is the control error signal distribution?

A. The control error signal distribution is a key factor in the calculation of the deployment costs. It describes the probability and magnitude of the one-minute error signal. The error signal represents the sum of the instantaneous deviation in load and the instantaneous departures in wind generation from schedule. These instantaneous departures are amounts of generation that the FCRPS must *inc* or *dec* in order to maintain load-resource balance in the BPA BAA during the operating hour.

Q. What amount of inc, dec, and spinning Operating Reserve were assumed for purposes of running the GARD model?

A. The average reserves need estimated for fiscal years (FY) 2010 and 2011 is 1,778 MW of *inc*, 2,347 MW of *dec*, and 256 MW of spinning Operating Reserve. These amounts were derived from the Generation Reserve Forecast and the Operating Reserve Cost Allocation. McManus *et al*., WP-10-E-BPA-23, section 4.7; and Study, WP-10-E-BPA-08, section 5, Table 5.3. These reserve quantities differ from those appearing in the Generation Reserve Forecast because Power Services (PS) began the process of studying the impact of reserves before the reserve forecast became final. The difference in reserve amount is 2 MW for the *inc* and 7 MW for the *dec*. These slight differences in reserve quantities do not have consequential impacts to the results of the Variable Cost Pricing Methodology Study and any difference between the reserve forecast and the variable cost methodology will be resolved in the Final Proposal.

SECTION 3: SPECIFIC VARIABLE COSTS CONSIDERED BY THE GARD MODEL

Q. What are the basic categories of variable costs used in the GARD model?

A. There are two broad categories: stand ready costs and deployment costs.

Q, Please describe the stand ready costs.

A. To meet the potential reserve requirements on any given hour, BPA's system must be set up to respond to these reserve needs going into the operational hour. Stand ready costs are those variable costs associated with ensuring that the FCRPS is operating in a way, at all times, that it is capable of providing the required reserve. In short, they are the costs that arise from ensuring that the FCRPS is standing ready to deploy reserves as needed. There are three sub-categories of stand ready costs: energy shift, efficiency loss, and base cycling losses.

Q. What are deployment costs and how do they differ from stand ready costs?

A. Deployment costs are those variable costs realized when the FCRPS actually deploys the standing ready reserve to automatically *inc* or *dec* generation in order to meet a within-hour reserve need that changes due to variations in loads and resources. The costs of meeting the within-hour variations in loads and resources are referred to as "deployment costs." There are four cost sub-categories for deployment costs: response losses, incremental cycling loss, incremental spill, and incremental efficiency loss.

Q. How does the act of providing reserves result in the type of stand ready cost quantified as energy shift cost?

A. Energy shift costs may be incurred while providing *inc* capability in circumstances where the ability to shape energy into the more valuable HLH period is limited due to lack of turbine availability. Turbine capability that could otherwise have been generating power for sale during the HLH period instead is rendered unavailable because turbines must be kept standing ready in case they are called on to increase generation for *inc* reserves. To the extent that providing the required *inc* capability is a contributing factor in limiting turbine availability, energy shifts into LLH.

Energy shift impacts also arise from the opposite situation, namely, making certain that sufficient *dec* capability (i.e., ability to decrease generation) exists during the LLH period (usually nighttime). In this instance, costs are incurred by taking energy from the HLH period and instead using it to generate during the LLH period to ensure that nighttime generation is sufficiently above minimum generation requirements to meet *dec* reserve needs.

Q. How are the energy shift costs calculated?

A. Energy shift costs are calculated by first determining how much energy is shifted out of the HLH period and into the LLH period in order to be standing ready to supply reserves. Second, the MWh amount shifted is multiplied by the difference between the HLH and

LLH prices from the market price forecast for the risk analysis. Study, WP-10-E-BPA-08, section 4.3.1.

Q. Why is the difference between HLH and LLH prices from the market price forecast for the risk analysis used?

A. HLH and LLH prices are used to capture the value difference between generating during the more valuable, high-demand HLH period versus the less-valuable LLH period. Energy shift costs are calculated by multiplying the amount of MWhs shifted from HLH to LLH by the difference in HLH and LLH price. The market price for risk analysis is used because if the energy was not shifted it would be available to market in the HLH, and this market price best captures the value of the shifted energy.

Q. How does the GARD model use the inputs described above to calculate the energy shift costs?

A. The GARD model uses the inputs to compare the HLH and LLH generation values for the no-reserves case to the test case where reserves must be carried. If energy is moved out of HLH and into LLH due to the need to increase LLH generation and/or graveyard generation and/or to ensure sufficient HLH *inc* capability, the HLH generation is reduced and the LLH generation is increased relative to the base case. Because HLH generation is more valuable than LLH generation, an economic impact is realized as HLH generation is reduced and LLH generation is increased.

Q. Does an energy shift cost always arise as a consequence of providing reserve?

A. No. To the extent the generation request plus the *inc* reserve is less than maximum generation, energy will not be shifted into the LLH period because there is sufficient generating capability available to meet the generation request and the full *inc* reserve.

Additionally, to the extent that the LLH generation is already above system minimum generation, there is no need to pull energy out of the HLH period. In these instances, "pre-existing *dec*" capability is said to exist. If the pre-existing *dec* capability

does not fully meet the *dec* requirement, energy is shifted out of the HLH and into the LLH.

Q. Is the energy shift cost dependent on the amount of spinning reserves versus non-spinning reserves?

A. No. Energy shift costs are the same whether the *inc* reserve is carried as spinning or non-spinning. The mere need for reserve capability is the cause of energy shift costs, regardless of whether that capability is available as spinning or non-spinning.

Q. In determining energy shift costs, why is the graveyard time period explicitly taken into account?

A. The graveyard hours (hours ending 0100 through 0400) are taken into account because the amount of pre-existing *dec* capability may be substantially less than what is available in hours ending 2300 through 0000 and hours ending 0500 through 0600. As explained above, it is necessary to maintain a cushion of generation above system minimum equal to the *dec* requirement to allow the FCRPS to decrease generation for balancing purposes. Because the graveyard hours have even less generation than ordinary LLH hours, there is an even greater need to shift generation into the graveyard to meet the system minimum necessary for the *dec* requirement.

Q. Based on the amount of reserves forecast in the Generation Reserve Forecast, what is the proposed forecast cost for energy shift?

A. Over the rate period of FY 2010-2011, the average energy taken out of the HLH period and shifted to LLH is 2,867,922 MWh, which results in a cost of $27,605,545.

Q. Are energy shift costs reflected in any other rates study?

A. Yes. The impact of the energy shift is implicit in the RiskMod studies. *See* Risk Analysis and Mitigation Study, WP-10-E-BPA-04, Operating Risk Analysis section. The RiskMod studies reflect energy shift costs as lost secondary revenues. The purpose of

calculating energy shift costs in the Variable Cost Pricing Methodology is to more appropriately allocate the lost secondary costs of providing reserves.

Q. Does RiskMod accurately reflect the energy shift cost?

A. RiskMod underestimates the energy shift cost, because it does not consider monthly time periods any more refined than the HLH and LLH blocks. As a result, an adjustment is passed to RiskMod based on the GARD graveyard adjustment to fully reflect the cost of the energy shift.

Q. Besides the adjustment for energy shift, are any of the other results from the Variable Cost Pricing Methodology provided to RiskMod?

A. Yes. The impacts of all remaining cost categories are provided to RiskMod, denominated in monthly HLH and LLH MWh amounts. The remaining cost categories, which we will explain in further detail below, include efficiency loss, base cycling losses, response losses, incremental cycling losses, incremental spill, and incremental efficiency losses.

Q. How are these MWh amounts taken into account in RiskMod?

A. The MWh impacts are accounted in RiskMod as a reduction in secondary sales revenue.

Q. Please describe the sub-category of stand ready costs known as efficiency loss.

A. Efficiency loss is the lost FCRPS power production capability associated with using available water less effectively because capacity is standing ready to provide reserves.

Q. Why does the act of providing reserves decrease the efficiency of the FCRPS?

A. Efficiency is, on average, decreased when standing ready to provide reserves because carrying reserves can result in an altered unit dispatch. For any given generation request, each project has an efficiency maximizing unit dispatch. If the most efficient unit dispatch results in an inadequate amount of reserve capability, the unit dispatch must be changed to stand ready to provide the required reserves. Because there exists one maximum efficient dispatch for any given level of generation, changing the unit dispatch for a given level of generation results in an efficiency loss.

Q. How does the GARD model use the inputs described above to calculate the efficiency losses?

A. The GARD model uses its inputs to calculate the change in efficiency under the base, no reserves, case as compared to the test case of providing reserves. For each month of each water year for both HLH and LLH, GARD finds the efficiency maximizing unit dispatch for GCL, CHJ, JDA, and TDA. The efficiency maximizing unit dispatch is found for the base case as well as the test case. Any observed difference in the efficiency between the base and test cases is the change in efficiency.

Q. Are efficiency losses always a consequence of providing reserves?

A. Changing the project dispatch may result in either an efficiency loss or gain; however, on average, altering the unit dispatch results in an efficiency loss. Gains may be achievable if the generation request inclusive of reserves allows the online units to operate closer to peak efficiency and/or the number of online units changes in a way that allows all online units to generate closer to peak efficiency. For example, under a high flow condition, carrying *inc* reserves may push units operating beyond peak efficiency back toward peak efficiency. Another example may occur on LLH, where increased *dec* reserve results in a higher generation level where additional units are brought online, potentially allowing for a more efficient generation dispatch.

Q. How does a spinning reserve obligation versus a non-spinning reserve obligation affect efficiency losses?

A. For any given *inc* reserve obligation, the greater the proportion of spinning capability, the greater the efficiency loss. Conversely, the greater the proportion of the reserve carried as non-spinning, the lesser the efficiency loss. This is true because the greater the spinning obligation, the greater the impact the obligation has on unit dispatch. Because spinning reserves are defined as unloaded turbine capability, increasing the spinning

obligation requires increasing the unloaded turbine capability. Unloading turbines generally results in decreasing the efficiency of the unit.

Conversely, the more *inc* reserve carried as non-spinning, the more the project is allowed to operate its spinning units more efficiently; it becomes less likely that turbines will have to be unloaded in order to meet the reserve obligation. The greater the non-spinning allowance, the more the *inc* obligation can be provided by idle turbines.

Q. How are efficiency losses priced?

A. Efficiency losses are priced using the market price forecast for the risk analysis HLH price. WP-10-E-BPA-03A, Table 18.

Q. Why is the HLH price used for all periods?

A. The HLH price is used for all periods because losses occurring during either the HLH period or the LLH period both result in decreased HLH generation. For the HLH period, HLH generation is directly lost during that period. For the LLH period, the same is true, because losses realized during the LLH period result in increased consumption of water that would otherwise have been used to generate during the HLH.

Q. Are efficiency gains also priced at the HLH price?

A. Yes.

Q. Are efficiency losses accounted for in any other rate studies?

A. No.

Q. Based on the amount of reserves described in the Generation Reserve Forecast, what is the proposed forecast cost for efficiency losses?

A. Over the rate period of FY 2010-2011, the average annual efficiency losses are 107,458 MWh for HLH and 179,432 MWh for LLH, resulting in an annual average cost of $15,352,534. Study, WP-10-E-BPA-08, section 4.3.2.

Q. Please describe the sub-category of stand ready costs known as base cycling losses.

A. Base cycling losses originate from the additional synchronization and ramping of units in order to ensure sufficient reserve capability is standing ready to respond.

Q. How are the base cycling losses calculated?

A. For base cycling, the number of units cycled on- or off-line is calculated by comparing the online units in the base case (i.e., no reserves) to the online units in the test case where the reserve requirement is being met. To the extent that more or fewer units are online when the reserve requirement is being met, a cycling cost is realized.

Q. How does the GARD model use the inputs described above to calculate the cycling loss?

A. The loss calculations are project-specific and are functions of the individual unit efficiency curves as well as the level of generation required from the individual units to meet the generation request for the amount of reserves required to stand ready. For each unit cycle, synchronization and ramping losses are calculated. During synchronization, water is lost as the unit is spun to synchronize to grid frequency. Water losses during synchronization are equal to 10 percent of full-gate-flow for three minutes. Ramping losses occur as the unit ramps up to its required generation level. Losses associated with ramping are calculated by evaluating the integral of the specific unit efficiency function from minimum generation to requested generation. The GARD model fully ramps units to their requested generation level over seven minutes.

Q. Why is the cycling loss assumed to occur on each HLH or LLH period?

A. Cycling losses are assumed to occur in each HLH and LLH period for a given month because the generation level for each HLH and LLH period in a given month is considered an average generation representative of all HLH and LLH periods in the month. That is, in the GARD model all HLH periods for a month are considered identical to each other, and all LLH periods for a given month are considered identical to each other. An observed unit cycle during any HLH or LLH period is said to occur for

each HLH or LLH period in the month. For example, if one additional unit is online during the HLH period relative to a case without a reserve requirement, that one additional unit is assumed to be needed for each HLH period in the month. That is, for a month with 18 HLH periods, 18 additional unit cycles would occur relative to the case without reserves.

Q. How are the base cycling losses priced?

A. All base cycling losses are priced at the monthly HLH price from the market price forecast for the risk analysis. WP-10-E-BPA-03A, Table 18.

Q. Why is the HLH price used for all periods?

A. The HLH price is used because the base cycling impact, that is, losses in energy, is taken out of the HLH period. In other words, because BPA seeks to maximize HLH sales, efficiency losses translate to lost HLH generation, with resulting lost sales.

Q. Does the GARD model or the Variable Cost Pricing Methodology include any additional maintenance costs associated with the additional cycling required by standing ready to provide reserves or deploying reserves?

A. The calculations of cycling losses and for that matter the incremental cycling cost, described below, do not attempt to account for any additional maintenance costs due to frequent cycling of the units. There are additional maintenance costs associated with the additional cycling of units to provide reserves, but we have not attempted to evaluate these additional costs, and all maintenance costs are included in the embedded cost analysis.

Q. Based on the amount of reserves described in the Generation Reserve Forecast, what is the proposed forecast for base cycling cost?

A. Over the rate period of FY 2010-2011, the average annual base cycling losses are 1,354 MWh for HLH and 2,572 MWh for LLH, resulting in an annual average cost of $214,154. Study, WP-10-E-BPA-08, section 4.3.3.

Q. Please describe deployment costs.

A. As discussed above, deployment costs are those variable costs realized when the FCRPS automatically *inc*s or *dec*s generation in order to balance the system. These are distinct costs from the standing ready cost. The cost sub-categories for deployment costs are response losses, incremental cycling loss, incremental spill, and incremental efficiency loss.

Q. Are there any overlaps between the deployment and the stand ready costs?

A. No. Deployment costs are realized only when units respond to a reserve need. Stand ready costs, on the other hand, are costs realized when setting up the FCRPS with the capability to respond.

Q. Please describe the sub-category of deployment costs known as response losses.

A. Response losses are a form of efficiency loss incurred when units online and on AGC respond to a signal. Response losses are an additional amount of efficiency loss realized as the unit's efficiency continuously changes over the course of deployment. The losses are a function of the respective controller project's unit dispatch, the project's response, and the amount of the control error signal.

Q. How are response losses different from efficiency losses?

A. Response losses are incurred when a plant that is standing ready to respond actually increases or decreases generation in response to a reserve need.

Q. How are response losses calculated in the GARD model?

A. The GARD model calculates response losses by simulating a control error signal and calculating how each of the controller projects' units change generation as a function of the given project's response allocation and the size of the error signal. When each of the units at a project experiences generation changes as a result of the simulated error signal, GARD calculates the average efficiency of the unit as it moves in response to the error signal. GARD calculates the average efficiency by integrating over the unit's efficiency

curve function from each unit's starting generation value to its ending value. The difference in the efficiency prior to deploying and the integrated efficiency during the course of response is the change in efficiency due to responding. Changes in efficiency are multiplied by the average generation during deployment to calculate losses in MWh, which are then totaled across generating units and controller projects.

Q. *Do the incs and decs balance each other over time?*

A. Yes. Over time, we assume that the energy consumed during the deployment of balancing reserves nets to zero.

Q. *If over time the expected average control error is 0 MW, why are response losses not zero on average?*

A. Response losses are not zero because the gains and losses in efficiency incurred during reserve deployment are not symmetrical. Stated another way, the change in efficiency for a one MW increase in generation is not equal to the change in efficiency for a one MW decrease in generation. The degree of this asymmetry is driven by the particular characteristics of a given unit as well as the unit's generation level and corresponding efficiency prior to reserve deployment.

Q. *How are response losses priced?*

A. Response losses are priced at the monthly HLH price from the market price forecast for the risk analysis. WP-10-E-BPA-03A, Table 18.

Q. *Why is the HLH price used for all periods?*

A. The HLH price is used because response impacts, losses and gains in energy, are taken out of or put into the HLH period.

Q. *Are response gains also priced at the HLH price?*

A. Yes.

Q. Based on the amount of reserves described in the Generation Reserve Forecast, what is the proposed forecast cost for response losses?

A. Over the rate period of FY 2010- 2011 the average, annual response loss is 31,397 MWh for HLH and 39,250 MWh for LHL, resulting in an annual average rate period cost of $3,922,246. WP-10-E-BPA-08, section 4.4.1.

Q. Please describe the sub-category of deployment cost known as incremental cycling losses.

A. During the course of deployment, an *inc* signal may exceed the available spinning capability. In these instances, the GARD model will synchronize and ramp additional units as needed to meet the *inc* signal. This process captures the effect of deploying non-spinning reserves. When additional units are brought online, cycling costs are realized.

Q. How are incremental cycling losses different from base cycling losses?

A. Base cycling losses are associated with the additional unit cycles required to meet a project's generation request and stand ready reserve requirement. Incremental cycling losses are experienced during reserve deployment as additional units are brought online.

Q. Is there overlap between response losses and incremental cycling losses?

A. No. Response losses apply to only those units that are already spinning when reserves are deployed. Incremental cycling losses are response losses associated with non-spinning units that are brought online to meet the *inc* signal.

Q. How are incremental cycling losses modeled?

A. Because the process of synchronizing and ramping takes place over 10 minutes, incremental cycles are modeled as occurring only once in any 10 minutes of the deployment simulation and only when an error signal exceeds the available spinning capability.

Q. How does the GARD model use the inputs described above to calculate the incremental cycling losses?

A. As spinning capability is exhausted, additional units are cycled until sufficient capability exists to respond to the control error signal. As with the previously described base cycling cost, the incremental cycling costs include synchronization and ramping losses.

Q. How are incremental cycling losses priced?

A. All incremental cycling losses are priced at the monthly HLH price from the market price forecast for the risk analysis. WP-10-E-BPA-03A, Table 18.

Q. Why is the HLH price used for all periods?

A. The HLH price is used because energy lost due to cycling is taken out of the HLH period.

Q. Based on the amount of reserves described in the Generation Reserve Forecast, what is the proposed forecast cost for incremental cycling losses?

A. Over the rate period of FY 2010-2011, the annual average incremental cycling loss is 15,553 MWh for HLH and 56,128 MWh for LLH, resulting in an annual average rate period cost of $3,923,586. Study, WP-10-E-BPA-08, section 4.4.2.

Q. Please explain why there is such a large difference in the proposed forecast between HLH and LLH.

A. The HLH period generally has more units online than the LLH period. A larger number of online units generally equates to a larger spinning reserve. As a result, for the same given control error, the LLH period is more likely to experience incremental cycling costs, because the spinning capability is more limited.

Q. Please describe the sub-category of deployment cost known as incremental spill.

A. Incremental spill may occur in the GARD model during the course of deployment one of two ways. The first is when a sufficiently large *dec* signal pushes generation below the amount of generation shifted out of the HLH and into the LLH. This occurs because the minimum generation constraint is driven by a flow rate requirement, so water must

continue to move past the projects at the same time the project is being required to reduce generation. The second occurrence of incremental spill is when the *dec* signal exceeds the project's maximum hydraulic ramp rate. To reduce generation more quickly, the project must spill continually to keep passing water while meeting the request to reduce generation.

Q. How is incremental spill calculated by the GARD model?

A. GARD watches for and calculates the impact of any incremental spill during the course of the error signal simulation. As discussed above, the error signal simulation models the need to deploy balancing reserves. For each minute of the error signal, GARD calculates how much it can decrease generation before needing to spill by comparing the *dec* error signal to the amount of generation shifted out of HLH and into LLH. To the extent that the error signal is less than the amount of shifted generation, no incremental spill occurs. If the error signal exceeds the amount of generation shifted into the LLH, the model relies on the pre-existing *dec* capability to meet the *dec* need. If the pre-existing *dec* is also exhausted before the *dec* need is met, the model spills as generation continues to be decremented if the minimum generation constraint is flow-based. This does not apply during periods where the minimum generation constraint is MW-based rather than flow-based.

Q. How does the GARD model account for unique characteristics of the Grand Coulee (GCL) project?

A. As stated above, spill may occur if the generation drop exceeds the drop rate allowed by the project. The drop rate constraint is a particular feature of GCL. GCL's ability to drop generation is limited because of tailwater bank stability concerns, which limit the rate of change in project outflow. More specifically, this limit is on how quickly the outflow may decrease. The tailwater constraint is determined by the United States Geological Survey (USGS) and enforced by the U.S. Bureau of Reclamation

(Reclamation). The tailwater constraint is represented in GARD as a function of GCL LLH generation.

Q. How is incremental spill priced?

A. All incremental spill is priced at the LLH market price from the market price forecast for the risk analysis. WP-10-E-BPA-03A, Table 18. This is because the spill is a consequence of relying on pre-existing *dec* capability, where pre-existing *dec* exists only when the system is already generating above LLH minimums and cannot shape the energy into HLH. In these instances, the system must continue to move the water on LLH even if generation is decreasing during a *dec* reserve deployment. Because the water must continue to move, the opportunity to shape into HLH does not exist, and only LLH generation is forgone.

Q. Based on the amount of reserves described in the Generation Reserve Forecast, what is the proposed forecast cost for incremental spill?

A. Over the rate period of FY 2010-2011, the average annual incremental spill for LLH is 181,778 MWh, resulting in an annual average rate period cost of $7,745,719. Study, WP-10-E-BPA-08, section 4.4.3.

Q. Please describe the sub-category of deployment costs known as incremental efficiency losses.

A. Incremental efficiency losses are losses that occur as the project attempts to efficiently dispatch in response to the error signal while ensuring that the spinning portion of the contingency reserve is maintained.

Q. How are incremental efficiency losses different from efficiency losses?

A. The incremental change in efficiency is distinct from stand ready efficiency loss in that it occurs during the deployment of reserves as the controller projects change dispatch as necessary to maintain their Operating Reserve obligation. As discussed above, the spinning Operating Reserve requirement must be maintained throughout a deployment.

These reserves cannot be used to respond to an *inc* signal caused by changes in load or generation unless the Operating Reserve is specifically called upon for a qualifying contingency

Q. How are incremental efficiency losses calculated?

A. Incremental efficiency losses are calculated by comparing the efficiency of each controller project at the beginning of reserve deployment to the efficiency after deploying. If during deployment the unit dispatch changed, the overall plant efficiency at the conclusion of the deployment is changed. After response, spinning units have been moved to a new generation level, and the response has potentially caused units to be cycled on/off line. The difference in the beginning efficiency and ending efficiency is the calculated change.

Q. How are incremental efficiency losses distinct from response losses?

A. The incremental efficiency loss is distinct from response losses in that incremental efficiency losses are the resulting efficiency after responding. In these measurements, the efficiency of the project after generation has changed to a new value in reaction to the error signal, while the response losses are associated with getting to the new generation level.

Q. How are incremental efficiency losses distinct from incremental cycling losses?

A. Cycling losses are realized during the synchronization and ramping of units during reserve deployment. The incremental efficiency loss is resultant plant efficiency after additional units have been brought online and is calculated at the conclusion of the reserve deployment.

Q. How are incremental efficiency losses priced?

A. All incremental efficiency losses and gains are priced at the HLH market price from the market price forecast for the risk analysis. WP-10-E-BPA-03A, Table 18.

Q. Why is the HLH price used for all periods?

A. The HLH price is used because efficiency impacts, that is, losses and gains in energy, are taken out of or put into the HLH period

Q. Based on the amount of reserves described in the Generation Reserve Forecast what is the proposed forecast cost for incremental efficiency losses?

A. Over the rate period of FY 2010-2011, the annual average incremental efficiency loss for HLH is 4,703 MWh, with an annual average efficiency gain of 14,749 MWh on LLH, resulting in an annual average rate period benefit of $543,022. Because this is a net gain in efficiency, it provides a credit to the calculation of the total variable cost of providing reserves.

Q. Why is there more efficiency gained than lost from the GARD model dispatch after a deployment of reserves?

A. Following the deployment of within-hour balancing reserves, the only reserve the model is required to continuously maintain during the hour is Operating Reserve. As a result, the model is often able to improve the efficiency of the projects for a given level of post-deployment generation relative to having been standing ready to deploy.

SECTION 4: APPORTIONING THE VARIABLE COST OF RESERVES

Q. What is the total forecast variable cost of providing reserves?

A. The total variable cost of providing reserves for the FY 2010-11 period is $58,221,062 annually based on the average amount of reserves described in the Generation Reserves Forecast and the spinning portion of the Operating Reserve described in the Operating Reserve Cost Allocation. WP-10-E-BPA-08, section 4.5.1.

Q. What is the purpose of apportioning cost to the different types of reserves?

A. The total variable cost of providing reserves is apportioned according to type of reserve to ensure cost recovery from the appropriate uses and causes. Specifically, the total cost

is apportioned into the cost of load Regulating Reserve, Load Following Reserve, Wind Balancing Reserve, and Spinning Operating Reserve. We allocate the costs of load Regulating Reserve, Wind Balancing Reserve, and Operating Reserve to transmission for TS to recover through Ancillary and Control Area Service rates. The remaining costs, those for Load Following reserve, are allocated to power for recovery through power rates.

Q. How is the reserve cost apportioned to reserve type?

A. The GARD model is run in a batch process where the costs of 25 different combinations of *inc* and *dec* reserve obligation are calculated. The batch run process is driven by the need to account for the cost diversity that exists when carrying different combinations of *inc* and *dec* reserves. The result of cost diversity is a lower cost for a given combination of *inc* and *dec* than the sum of the individual costs for *inc* alone and *dec* alone. The batch model run is the first step in determining a diversified cost separation.

Q. Please describe what is done with the results of the batch run.

A. The costs obtained from the batch model run are broken into spinning and non-spinning costs. Spinning costs are assigned the energy shift cost associated with the spinning *inc* obligation and the *dec* obligation, the base cycling cost, efficiency losses, and response losses. Each of these cost categories is associated with units online and generating. Non-spinning costs are assigned the energy shift cost associated with the non-spinning portion of the *inc* obligation, incremental cycling losses, incremental spill, and incremental efficiency losses. Each of these costs is realized as units are cycled on from non-spinning status or cycled off to non-spinning status. These spinning and non-spinning costs are shown in the Study, WP-10-BPA-E-08, Tables 4.6 and 4.7.

Q. How are the tables of spinning and non-spinning costs used?

A. The tables of spinning and non-spinning costs are used to generate parameter estimates for two multivariate linear regressions. One regression describes spinning reserve costs

as a function of the spinning *inc* and total *dec* obligation. The second regression describes non-spinning reserve costs as a function of non-spinning *inc* and total *dec* obligation. The multivariate linear regression is a mathematical model used to describe the reserve costs as explicitly calculated by GARD as a function of multiple variables; in this case *inc*s and *dec*s. Stated differently, the equation is used to calculate reserve costs using the *inc* and *dec* obligation as inputs. The results from the equation closely mimic those results explicitly generated by the GARD model. With the equations, total costs are now the sum of the spinning and non-spinning costs for a given *inc* and *dec* combination. The linear regression is used to allocate costs by reserve type. WP-10-E-BPA-08, sections 4.5.3 and 4.5.4.

Q. Why are non-spinning costs partially a function of decs?

A. The *dec* obligation is used as an explanatory variable, that is, an input in a mathematical function, for non-spinning costs, because costs from cycling units offline and/or spilling while deploying to meet a *dec*, as well as the resulting plant efficiency changes, are all rolled into non-spinning costs. Putting a unit into non-spinning status during a *dec* deployment is the opposite of bringing up a unit from non-spinning status during an *inc* deployment. Thus, it is appropriate to use the *dec* obligation as the explanatory variable for both situations.

Q. Why are the component costs calculated as relative to the total cost?

A. Component costs are calculated relative to the total cost as calculated by the regression model. Although the regression model fits the cost data as calculated by GARD very well, the fit is not perfect. The result of the imperfect fit is a slightly different total cost for a given reserve as calculated using the regression equations compared to the total cost GARD calculates. In order to eliminate the impact of the differences between the regression-based calculation and the GARD model run, the cost for each reserve category as derived from the regression is taken relative to the total cost as calculated by the

regression. These relative costs per reserve category are then applied to the total cost GARD calculates to arrive at a dollar cost per reserve category. WP-10-E-BPA-08, section 4.5.5.

Q. In reviewing the GARD model after the calculations used in the Initial Proposal were completed, did you discover any discrepancies in the model?

A. Yes. An error in the code caused the model to use the incorrect synchronization and ramping cost during large *dec* events when units were cycled off. The model incorrectly used the cycling loss from a previous iteration's *inc* event. The impact of correcting this would be an increase in the incremental cycling costs for FY 2010-2011 from an annual average cost of $3,923,586 to an annual average cost of $5,244,666.

Q. What effect do you anticipate this correction would have on the results of variable cost allocation forecast?

A. The correction increases the total variable cost allocation forecast by approximately 2.3 percent, raising the total average annual cost for FY 2010-2011 from $58,221,062 to $59,542,141. This change adds an estimated $79,154 to Regulating Reserve and $1,025,508 to Wind Balancing Reserve, and reduces Operating Reserve costs by $281,870 compared to the results reflected in the Initial Proposal. Additionally, the Load Following cost in power rates is estimated to increase by $498,288 compared to the Initial Proposal.

Q. Why does this correction decrease the variable cost allocated to Operating Reserve?

A. The difference is due to the apportioning of costs being performed on a relative basis. Slight differences in the percent of total cost allocated to Operating Reserve result in slightly different result when the ratio is applied to the total cost GARD calculates. Although the batch simulation produced the same absolute cost for Operating Reserves, the relative cost changed slightly.

Q. Why is this correction not reflected in the Initial Proposal?

A. This correction is not reflected in the Initial Proposal because there was insufficient time due to publication deadlines to fully incorporate the results into all processes and calculations affected by the results of the Variable Cost Pricing Methodology.

Q. How will you incorporate this correction in the Final Studies?

A. The GARD runs for the final Generation Inputs Study will have this correction integrated into all processes reliant on the results of the Variable Cost Pricing Methodology.

SECTION 5: IMPACTS OF CHANGING THE PERSISTENCE SCHEDULING ASSUMPTIONS ON THE VARIABLE COST ALLOCATION

Q. Were any other variable cost forecasts generated?

A. Yes. In addition to the studies performed for the specific reserve need of TS based on the two-hour persistence forecast described in the Generation Reserve Forecast, cost analysis was performed for three additional scenarios. Each scenario assumes increasing scheduling accuracy on the part of the wind fleet contained within the BPA BAA. Scheduling accuracies equivalent to 60-, 45-, and 30-minute persistence forecasting were analyzed. For a description of persistence accuracy assumptions, *see* McManus *et al*., WP-10-E-BPA-23, section 4.4 and 5. The amount of the *inc* and *dec* obligation for each scenario is taken from the Generation Reserve Forecast, Table 2.11, and re-run through the GARD model in the same fashion as the base case, which relied on 2-hour persistence derived values.

Q. What is the total forecast variable cost of providing reserves given the 60-, 45-, and 30-minute studies?

A. The total annual variable cost of providing reserves for the fiscal year 2010-11 periods is $50,521,870 for the 60-minute study, $42,602,876 for the 45-minute study, and $35,985,663 for the 30-minute study. These totals are for providing the average amount of reserves described in the Generation Reserve Forecast, Study, WP-10-E-BPA-08,

Table 2.11, and the spinning portion of the Operating Reserves described in the Operating Reserve Cost Allocation, Study, WP-10-E-BPA-08, Table 5.3. The results of these different persistence scheduling assumptions for the specific component costs are described in the Variable Cost Pricing Methodology, Study, WP-10-E-BPA-E-08, section 4.6 and tables 4.20, 4.21, and 4.22.

Q. Does this conclude your testimony?

A. Yes.