Spreadsheet Model for Benchmarking Ontario Power Distributors

User's Guide

Dave Hovde

Pacific Economics Group Research LLC

May 2015

Contents

This user's guide is designed to help local distribution companies ("LDCs" or "distributors") better understand the cost benchmarking methodology, concepts and calculations approved by the Ontario Energy Board ("OEB") for Ontario Power Distributors. OEB staff engaged Pacific Economics Group Research ("PEG") to develop an enhanced benchmarking Spreadsheet Model and a User's Guide for electricity distributors in relation to the OEB's Implementation of Improvement Initiatives for the 2014 Electricity Distributor Scorecard. PEG was assisted by members of an industry Scorecard Working Group, a Benchmarking Process Improvements Subcommittee and OEB staff from February to April 2015 in developing these documents.

The calculations are done in a Microsoft Excel spreadsheet titled "Spreadsheet Model Benchmarking Ontario Power Distributors." The spreadsheet performs the calculations necessary to arrive at a cost benchmarking result for 2013 and subsequent years under the IRM-4. A validation worksheet has been added to increase the transparency of the benchmarking results so that a distributor can readily review its benchmarking results, easily trace the source and use of data, and review the various calculations including distributor-specific parameter estimates. Validation steps are provided to aid distributors in the review of their benchmarking results. It also provides a forecasting capability to allow distributors to estimate their future benchmarking performance. Benchmarking results for earlier years were calculated in complex computer code as part of the econometric work undertaken as part of IRM-4. Calculations for earlier years may be included in the spreadsheet model at a future date.

The guide consists of several sections. The first section provides some background on the cost benchmarking methodology taken from previous PEG reports. The second section provides a broad overview of the spreadsheet model. The third section explains the calculations involved in the calculation of total cost. The fourth section discusses the econometric model and the calculations involved with obtaining predicted cost and the cost performance evaluation. This is the most difficult part of the calculations. Section five discusses the forecasting capability of the workbook. Section six reviews a flowchart included to upgrade the presentation of the 2013 working papers. Section seven presents generic calculation worksheet included as a learning tool. The last section discusses procedures for correcting errors and possible future improvements.

Note that while this user's guide was developed using specific references to the contents of the information in the 2013 "Spreadsheet Model Benchmarking Ontario Power Distributors", the guide is intended to serve as a tool for distributors to the end of the IRM 4 term. To facilitate the use of this guide during the IRM 4 period, all subsequent years' Spreadsheet Model Benchmarking Ontario Power Distributors will have consistent worksheets, formats and cell references to that of the 2013 Spreadsheet

Model. As such, the guide can be used interchangeably with the subsequent years' models by substituting the applicable year in use at the time.

1 Background

1.1 Procedural History

In 2013, as part of the IRM-4 proceeding EB-2010-0379, the OEB issued a report titled "Rate Setting Parameters and Benchmarking under the Renewed Regulatory Framework for Ontario's Electricity Distributors"[1] ("Board Report") in which it set forth the framework for setting rate adjustment formulas for LDCs. Links to key benchmarking documents are provided in Appendix 1. The Report provides the OEB's final determination on its policies and approaches to the distributor rate adjustment parameters and the benchmarking of electricity distributor total cost performance for the 2014 to 2018 rate period. In addition and consistent with these policies and approaches, the Report sets out the inflation factor and the stretch factors to be used in 2014 incentive rate setting (i.e., the Price Cap Incentive Rate-setting and the Annual Incentive Rate-setting Index) applications.

According to the Board Report, rates will be indexed by a formula "which is used to adjust the distribution rates to reflect expected growth in the distributors' input prices (the inflation factor) less allowance for appropriate rates of productivity and efficiency gains (the X-factor)."[2] The productivity part of the X-Factor is based on the industry trend and is the same for all LDCs[3]. The efficiency gains part of the X-Factor is called the stretch factor and can vary by company. This stretch factor reflects the potential for incremental productivity gains by a given LDC under incentive regulation which in turn depends on an individual distributor's level of cost efficiency. Additional discussion of the rate adjustment formula is contained in Appendix 2.

The stretch factor assignments are based on the results of a statistical cost benchmarking study designed to make inferences on individual distributors' cost efficiency. An econometric model is used to predict the level of cost associated with each distributor's operating conditions. Distributors that had actual cost that was lower than that predicted by the model were assigned lower stretch factors than those

[1] Issued on November 21, 2013 and corrected on December 4, 2013.

[2] Board Report, page 5.

[3] The Board's approved productivity factor of 0% was informed by a study of the total factor productivity trend of the industry. Hydro One and Toronto Hydro were excluded because their large size would dominate the calculation of the trend that was to be applied to all distributors.

that did not. The October 18, 2013 report by Pacific Economics Group study ("PEG Report") titled "Productivity and Benchmarking Research in Support of Incentive Rate Setting in Ontario" describes the econometric model used to produce the benchmarking results. This work was updated in 2014 to include cost performance measures for 2013. The report titled "Empirical Research in Support of Incentive Rate-Setting: 2013 Benchmarking Update" discusses this update.

1.2 Evolution of Cost Benchmarking in Ontario

The preferred cost benchmarking method has evolved over time. In the era following IRM-3, only OM&A cost was benchmarked. During the IRM-4 proceeding, the scope of benchmarking was expanded to include capital cost as well. Starting in 2013, the cost benchmarking results were used to set 2014 stretch factors based on performance through 2012. The current method for assigning stretch factors was adopted at that time. It consists of five cohorts with fixed boundaries. Those distributors showing cost performance within certain ranges were assigned the same stretch factor. The ranges for each cohort are as follows:

Cohort I (0.00%): Cost < 25% of Predicted

Cohort II (0.15%): Cost Between -10% and -25% of Predicted

Cohort III (0.30%): Cost Between +10% and -10% Predicted

Cohort IV (0.45%): Cost Between +10% and +25% Predicted

Cohort V (0.60%): Cost > 25% of Predicted

The five cohort system evolved from a three cohort system to allow for greater mobility between cohorts and therefore better incentives for cost control. The cohort system also has the benefit of allowing for small data changes and errors to be corrected without affecting the assigned stretch factor. There have been numerous changes to data and corrections of the calculations which have not yet resulted in a different stretch factor assignment for a distributor.

1.3 Econometric Model

Distributor cost in the econometric model is estimated as a function of select business conditions faced by each distributor. These business conditions include the number of customers served and the price of inputs such as labor and capital. The business conditions included reflect those that were identified as having a statistically significant impact on cost by the model. Some business conditions such as the amount of high voltage transformation were explicitly taken into account by adjusting the definition of cost to be benchmarked. Other business conditions such as the area of service territory and percentage of line that is underground were tested and reviewed but not found to be statistically significant in their impact on cost for Ontario power distributors. As such, while not all business conditions faced by a distributor were included in the model, those selected by the model have statistically significant relationship to cost. The parameters of this model establish the relationship between each selected business condition and distributor cost. These parameters were estimated using Ontario LDC data from 2002-2012 as part of the IRM-4 proceeding. The model therefore provides the relationship between cost and business conditions faced by a typical power distributor in Ontario.

The econometric model can make a prediction of each distributor's cost given identified business conditions by multiplying the values for the company's business condition variables by the model parameters and summing the results[4]. The distributor's actual cost is then compared to that predicted by the model. The percentage difference between actual and predicted cost is the measure of cost performance. When actual cost turns out to be lower than the predicted cost, this demonstrates a better cost performance. As such, companies with larger negative differences between actual and predicted costs are considered to be better cost performers and therefore eligible for lower stretch factor classifications (see section 1.2 above).

For example, suppose that the cost for a power distributor could be described by this simple equation: Cost = $1,000,000 + $300 x Number of Customers. A distributor with 1,000 customers would have cost of 1.3 million dollars that is predicted by this equation. If the distributor on the other hand with 1,000 customers had a cost incurred that was actually only 1 million dollars, this amount would be compared to that predicted by the model and the conclusion would be that cost performance was better than expected by about $300,000. This amount translated into percentage terms would be their cost performance measure.

An econometric model uses the statistical technique of regression analysis to specify the parameters of an equation that best fits the observed data. In the context of power distribution cost, the parameters of a cost function were estimated[5]. In the simple equation Cost = A + B x Output, both cost

[4] The table of parameters published in the PEG report was for the full sample. When making predictions of cost for each company, the econometric program estimated the model without including the subject of benchmarking in the sample. Therefore, there exist 73 different sets of parameters which are very similar to each other. For ease of presentation, the PEG report did not present the parameters specific to each distributor. These company-specific parameters are necessary for the 2013 calculations and are contained within the working papers associated with this report.

[5] The form of the equation was a "translog" cost function which has many desirable properties which include the ability to provide an approximation to an arbitrary underlying production technology. This

and output are known while A and B are the unknown parameters. The regression method is designed to find the combination of the unknown parameters that best fit the observed data. In the simple example above, the values of $1,000,000 and $300 were the values for the parameters. For each observation, the parameters times the explanatory variables will generate an estimated value for cost. The difference between this estimate and the actual value is called the error. The square of this error will measure positive and negative deviations from the actual value equally and treat larger deviations from the actual value as being more serious than smaller deviations. The criteria for “best fit” is to minimize the sum of this squared error over all observations.

The estimated parameters establish the relationship between cost and each explanatory variable. Once estimated parameter values are available, it becomes possible to obtain a predicted cost value for any combination of explanatory variables including future values for a company that were not included in the original estimation of the model. The difference between the cost predicted by the model using the observed explanatory variables for a distributor and the actual cost of the distributor is the basis for the cost performance evaluation. A detailed description of the econometric model including estimation technique and other technical details are contained in sections 6 and A2.1 of the PEG report and are included as Appendix 3 to this guide.

The econometric model used to obtain the updated stretch factors is identical to the model described in the PEG report. The Board intentionally decided not to update the parameters of the econometric model to include 2013 data. The goal was to establish a fixed benchmark that would allow companies a fair opportunity to demonstrate improved cost performance and earn a lower stretch factor. The rationale for this decision is discussed in the Board Report and in a memorandum by PEG that also makes some corrections to the 2012 results.[6] The PEG memorandum contains the corrected final results of the 2010-2012 benchmarking model used in this update.

2 Overview of the Spreadsheet Model

The Spreadsheet model contains many worksheets. These were necessary for the simultaneous calculation of benchmarking results for all Ontario LDCs. The worksheet titled “2013 Benchmarking Calculations” contains the calculations used to generate the benchmarking results for all distributors included in the study period. A review of these benchmarking calculations by a distributor is presumably

helps avoid debate about the correct form of the equation when the exact relationship between cost and determinates of cost is not known.

[6] Available on the OEB website in the file “PEG_Memorandum_OEB on_corrections_20131220.pdf”+

concerned only with its own benchmarking calculations. It is therefore useful to create a single "validation" sheet that performs the identical calculations for a selected company on a distributor-specific basis. This increases the transparency of the benchmarking results in that it ensures a distributor can readily review its benchmarking results, easily trace the source and use of data, and review the various calculations including distributor-specific parameter estimates and values (discussed in detail below). . Future versions may include calculations for previous years.

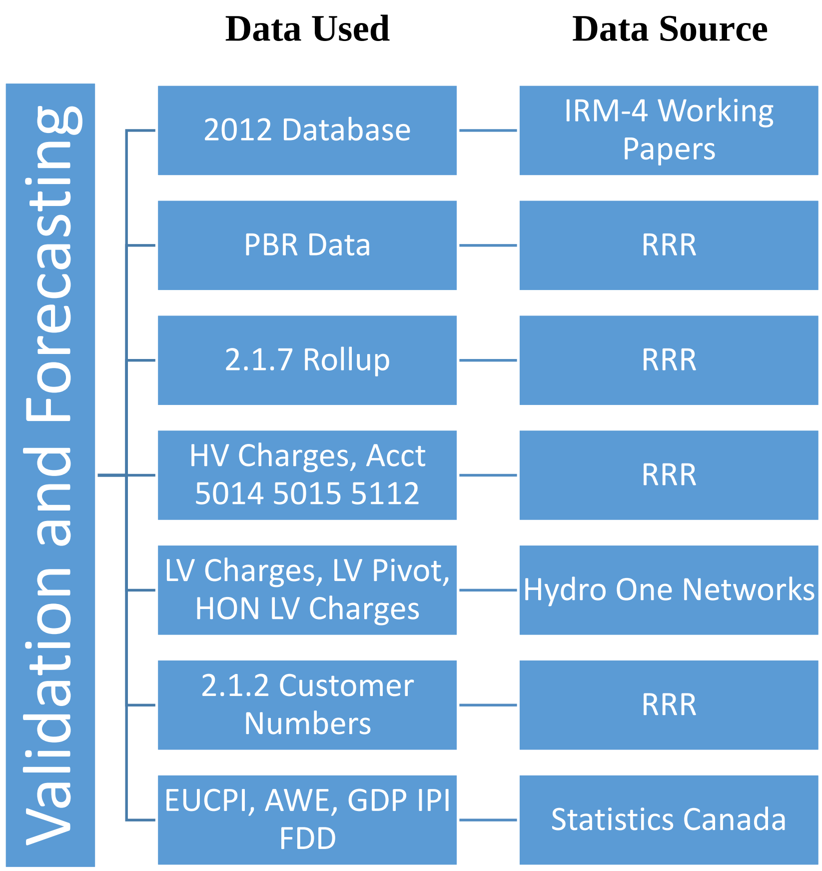
The worksheet titled "Validation" contains all the data and formulas required for the validation of the 2013 cost performance for a single distributor. It will refer to the other sheets to gather the necessary data, but all the formulas are on this sheet. It is equivalent to the 2013 worksheet and contains the same formulas and will produce the same results. The difference is that is has some additional OM&A detail and is formatted to make it easier for a distributor to review. Figure 1 shows the relationship of the other worksheets to the Validation worksheet and the source of the data.

With the exception of referencing the parameter values, the validation worksheet does not rely upon the 2013 calculations worksheet. The file will be revised in the future to put the parameter values on a separate sheet that both calculation sheets can reference independently. The remainder of the discussion below will focus on the Validation worksheet.

The validation worksheet also contains formulas for future years to allow distributors to predict future benchmarking results. It is designed to provide results for any single LDC. Cell E3 of the worksheet contains the name of the currently selected LDC. Clicking on this cell will display a drop down box where any other LDC may be selected. The worksheet will then update to reflect the data for the new LDC.

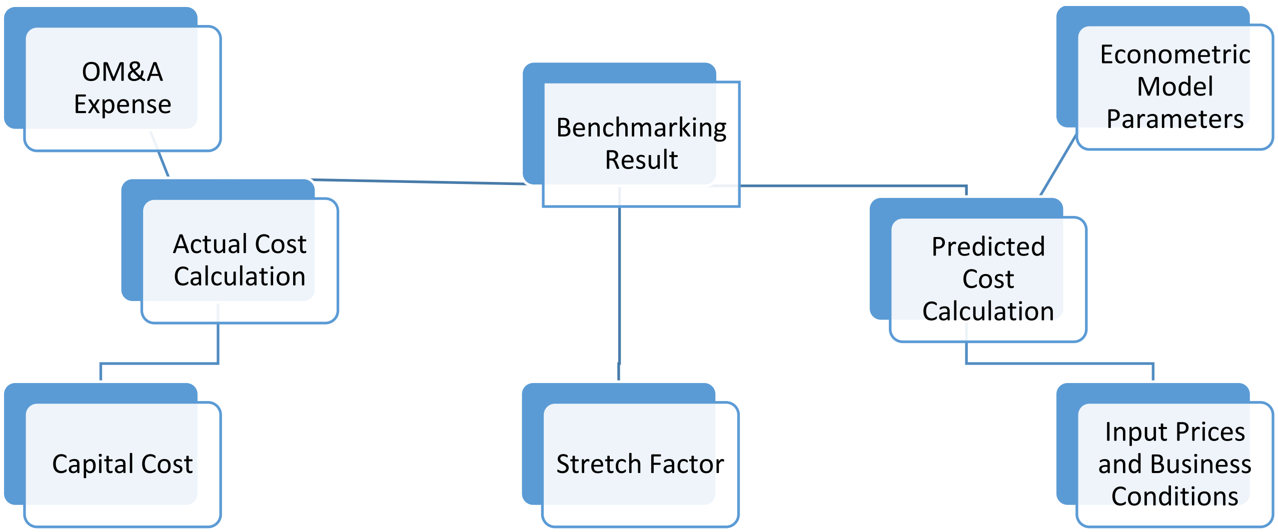
The goal of the calculations is to obtain an evaluation of an LDC's cost performance. This is done by comparing the total cost for the LDC to a prediction of total cost by an econometric model. The econometric model can make a prediction of each distributor's cost given its values for the business conditions by multiplying the company's business condition values by the model parameters and summing the results. The distributor's actual cost is then compared to that predicted by the model. The percentage difference between actual and predicted cost is the measure of cost performance.

Figure 1



A large number of calculations are required to obtain these benchmarking results and at times the formulas will be complex. In order to facilitate review, it is useful to break down the calculations into pieces. The first logical division of the calculations is the calculations involved in the calculation of actual cost vs. those for predicted cost. Figure 2 shows how these calculations may be organized.

Figure 2



The term "actual cost" is being used here to distinguish it from predicted cost and it is not intended to imply that this definition of cost is superior to those definitions used for non-benchmarking purposes. The OM&A definition has been chosen to allow better comparisons among distributors. The capital cost calculations use methods that are standard practice in productivity and cost research. The details of the Actual and Predicted Cost calculations are discussed in sections three and four respectively.

3 Calculation of Actual Cost

3.1 Background and Theory of the Actual Cost Calculation

The actual cost calculations can be broken down into calculations for OM&A and capital. The definition of cost was chosen in order to facilitate a fair comparison among companies and include some adjustments to allow for a level playing field. Examples include the addition or removal of certain cost items. For OM&A, high voltage costs were removed to control for the different scope of work done by LDCs. Because low voltage charges from Hydro One Networks to LDC are not captured by the OM&A accounts used for benchmarking, these charges are added.

Capital cost is calculated as the product of a price and quantity. The quantity is calculated using the perpetual inventory method. Each year, an amount of capital is discarded based on the depreciation rate and new capital is added based on gross plant additions. These amounts are quantities in the sense that price inflation has been removed. This method is preferred for benchmarking because it will help control for differences in the vintage of plant. A cost comparison based on the historical cost of plant would be biased. All else equal, a company with older plant could be seen as a superior cost performer because its plant and equipment were acquired at a time when construction costs were lower. The capital cost method is described in the PEG report provided as part of IRM-4[7]. The capital price equation takes into account the rate of return and depreciation. Both of these were standardized such that differences in the allowed rate of return, depreciation practices, and when debt was incurred would not influence measured performance.

One consequence of the perpetual inventory method is that calculations from previous years carry forward into the current year calculations. For 2013, the quantity of capital for 2012 is used, but not independently calculated in the spreadsheet. The support for this value is found within the working papers provided for IRM-4. A link to these working papers and associated documentation can be found in Appendix 1. The integration of the IRM-4 working papers into the spreadsheet model was not considered a priority at this time, but could be undertaken at a later date.

The figures below contain the relationship between the various data items. The numbers in parentheses refer to the line reference numbers contained in column B of the Validation worksheet. The flowcharts below are designed to start with actual cost and show what data items are used to arrive at this result. The calculations on the spreadsheet will start at the bottom of the flowchart and eventually arrive at the result at the top.

---

[7] The capital cost method standardizes the cost of capital such that it estimates how much it would cost a distributor to "rent" the quantity of capital that it actually owns. In doing so, the quantity of capital for each company is calculated in the same manner and is depreciated in a manner that mimics the market for used assets. The capital price translates this quantity into current dollars and then applies a rate of return and a depreciation rate that is the same for all companies. It calculates rate of return and depreciation measures that may appear similar to those used in ratemaking, but they are intended to represent how much the owner of the capital would need to charge a distributor leasing these assets in order to cover their cost in a competitive market. It is not intended to mimic the carrying cost of these assets in a regulated environment which will heavily depend on factors such as when the plant was constructed.

Figure 3

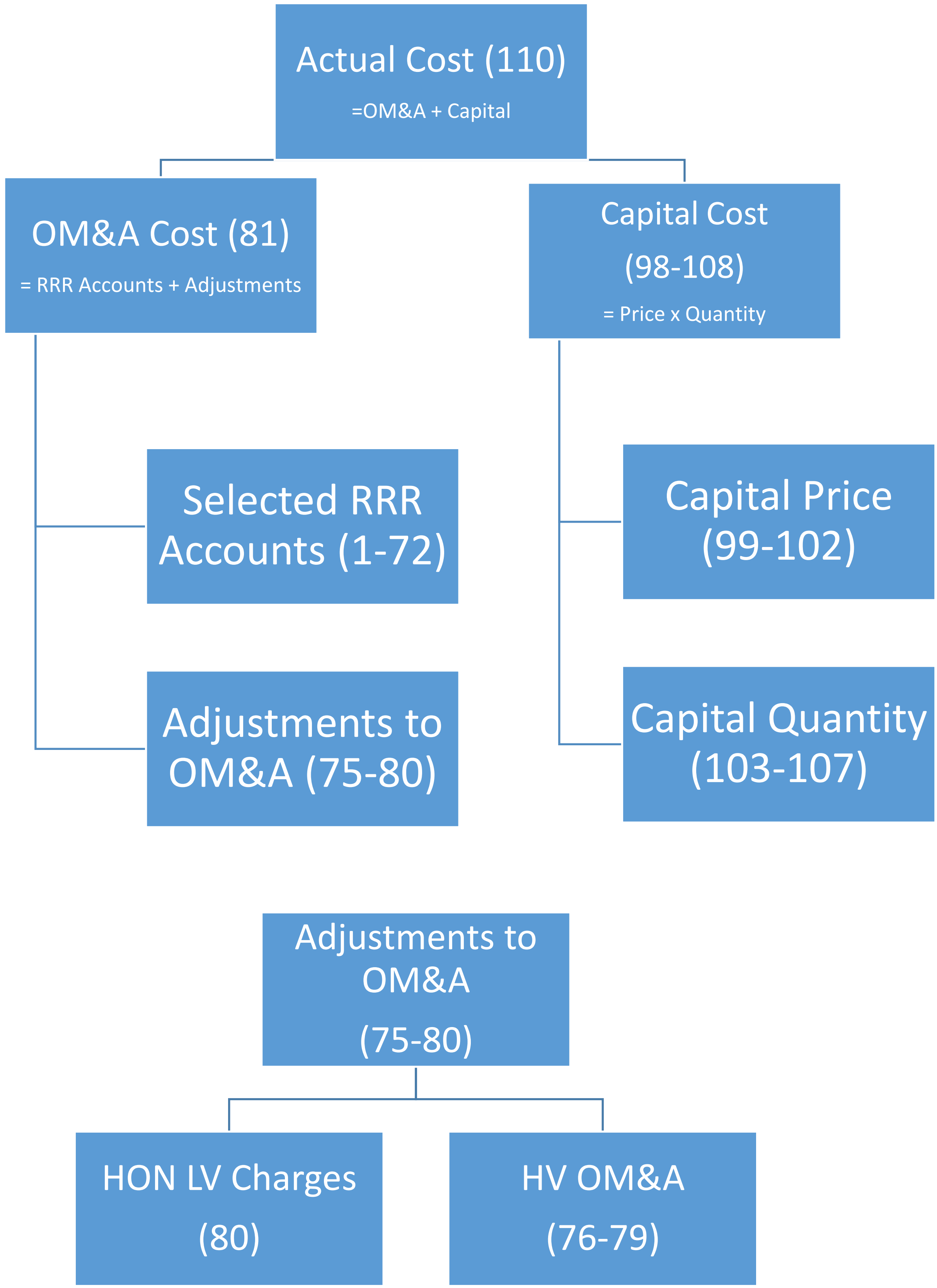
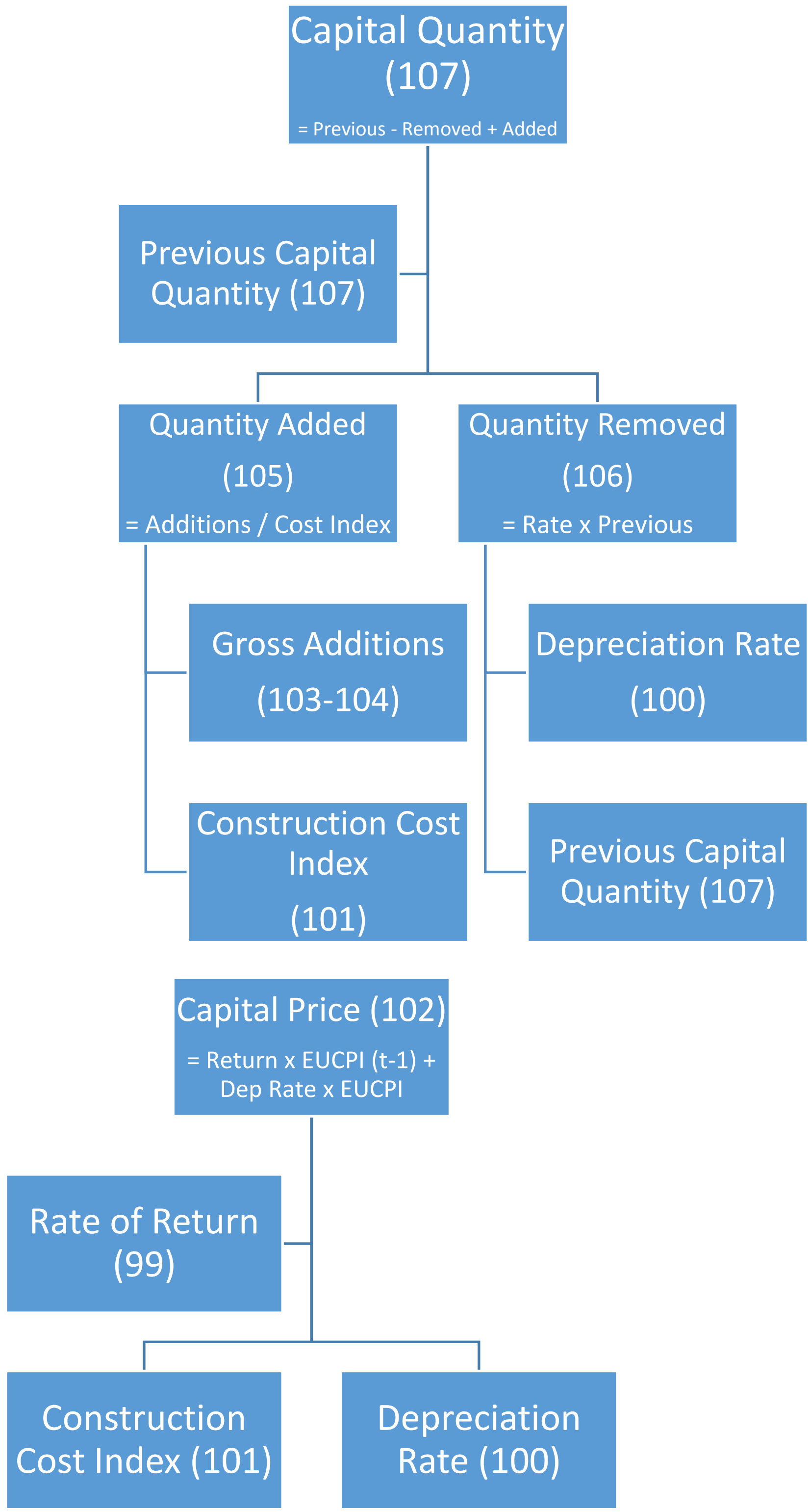

Actual Cost (110)
=OM&A + Capital
OM&A Cost (81)
= RRR Accounts + Adjustments
Capital Cost
(98-108)
= Price x Quantity
Selected RRR Accounts (1-72)
Adjustments to OM&A (75-80)
Capital Price (99-102)
Capital Quantity (103-107)
Adjustments to OM&A
(75-80)
HON LV Charges
(80)
HV OM&A
(76-79)

Figure 4


Capital Quantity (107)
= Previous - Removed + Added
Previous Capital Quantity (107)
Quantity Added (105)
= Additions / Cost Index
Quantity Removed (106)
= Rate x Previous
Gross Additions (103-104)
Depreciation Rate (100)
Construction Cost Index (101)
Previous Capital Quantity (107)
Capital Price (102)
= Return x EUCPI (t-1) + Dep Rate x EUCPI
Rate of Return (99)
Construction Cost Index (101)
Depreciation Rate (100)

The calculations on the spreadsheets are discussed below. Reference lines 1 through 110 will be covered here and the remainder in the next section.

3.2 Review of the Actual Cost Calculations on the Validation Worksheet

Section 1 of the Validation worksheet gathers most of the data required for the calculations from other sheets in the workbook. It also calculates the OM&A measure applicable to benchmarking. Lines 1 through 72 gather account level detail from the RRR 2.1.7 rollup worksheet which contains OM&A data. Subtotals are calculated to conform to previous groupings of OM&A cost. To validate these results, the LDC should verify that the each of the values for the listed accounts correspond to the data provided to the Board. An examination of each of the subtotal formulas on lines 22, 36, 44, 49, 67, 70 and 72 should show that these properly add up the included accounts.

Total OM&A benchmarking cost prior to adjustments is calculated on line 73 as the sum of the subtotals calculated above. It should be noted here that the definition of OM&A cost used for benchmarking will differ from that commonly used by the distributor. It excludes items such as bad debt expenses and most CDM cost (i.e. funded outside of distribution rates) to facilitate fair comparisons among distributors. It also excludes items such as amortization and taxes. Amortization has been standardized and already built into the capital cost methodology and is therefore not included. Taxes are excluded because they are largely outside of management control. To validate this result, an examination of the formula on line 73 should show that this is the sum of the subtotals previously validated.

Adjustments to OM&A are gathered on lines 75 through 80. These adjustments are for HV and LV cost. The HV adjustment removes the cost associated with the three RRR accounts related to HV service. The LV adjustment adds OM&A cost not included above associated with purchases of LV service from Hydro One Networks. These cells reference other sheets later in the workbook. In the case of LV service, a formula approved by the IRM-4 working group was used to identify the specific accounts and portion of accounts that was to be used for benchmarking. The original file from Hydro One is included along with the formula used to add up the selected accounts. The formula used is available on line 23 of the LV Charges worksheet[8]. Because not all companies are embedded in Hydro One and thus not represented in the Hydro One file, the data were arranged such that companies without data generated zero values.

[8] The formula sums Meter Charges (2C), Specific ST Lines (3B), LVDS (3D), and 45% of HVDS Low (3G)

Total OM&A expenses applicable to benchmarking are calculated on line 81 as the sum of the previous total on line 73 plus the two adjustments. The remainder of the distributor data required for benchmarking is gathered on lines 83 through 91. Lines 83 to 91 contain LDC-provided data on gross additions, output and other business conditions used in the analysis. To validate, these values should be checked against what was reported to the Board. The following is a summary of the suggested validation steps for Section 1:

1. Confirm that the values for each account listed match what was reported to the Board
2. Confirm that the subtotals of these items on 22, 36, 44, 49, 67, 70 and 72 correctly sum these accounts
3. Confirm that the formula on line 73 sums the subtotals
4. Confirm the values for the HV accounts 5014, 5015, and 5112 and the total on line 79
5. Confirm the billing data provided by Hydro One Networks (if any) and review the formula to calculate LV charges from the billing data
6. Confirm the formula on line 81 that calculates the OM&A expense used in benchmarking
7. Confirm that the gross additions, output, and other business condition data provided on lines 83 to 91 conform to that reported to the Board.

At this point, OM&A cost should be validated as well as for all LDC provided data required for the analysis.

Section 2 of the worksheet calculates capital cost and actual total cost. The capital cost calculations are done on lines 98 through 108. The capital price calculations are done on lines 98 through 102. Weighted average cost of capital (WACC) information was provided by Board Staff and includes changes in the Board approved ROE and interests rates on debt. Because this information does not conform to the calendar year, the WACC using in the benchmarking work was calculated as a weighted average that reflects how many months each WACC in effect during the year. The depreciation rate was determined in the IRM-4 work. It was calculated as a weighted average of the economic depreciation rates for several types of plant. These deprecation rates are geometric and a direct comparison with straight-line rates is not possible. The construction cost index on line 101 escalates the construction cost index value used in IRM-4 by the growth in the Electric Utility Construction Price Index as published by Statistics Canada. Because this index is subject to revisions and rebasing, the previously used values are escalated instead of using the index number level. This will prevent revisions from previous years from influencing current year benchmarking results. The capital price index is equal to the annual amortization of plant plus the rate of return on one unit of capital.

Capital quantity and capital cost are calculated on lines 103 through 108. The capital quantity calculation on line 107 starts with the previous capital quantity from the IRM-4 work. It reduces this quantity by the amount depreciated as calculated on line 106 and then adds the quantity of new capital calculated on line 105. The quantity of capital added is equal to the cost of the additions divided by the price of additions given by the construction cost index. The cost of capital calculated on line 108 is the product of the capital price and capital quantity. The total cost on line 110 is the sum of capital cost and OM&A cost.

To validate the capital cost and total cost results, the following steps are recommended:

1. Note that the rate of return and depreciation rates on lines 99 and 100 are values and therefore will be the same for each LDC.
2. Note that the 2012 values for the construction price index and capital quantity index on lines 101 and 107 are required. Also note that these values are taken from the IRM-4 worksheets provided.
3. Review the Construction Cost Index calculation on line 101. Observe that it escalates the value used in the IRM-4 work.
4. Review the capital price formula on line 102. Note that it contains rate of return and depreciation parts that are standardized substitutes for these items used in standard utility accounting.
5. Observe that the formulas in lines 103 and 104 “pull down” the gross additions data previously validated in Section 1.
6. Review the formula for Quantity of Capital Additions on line 105. Note that it removes HV additions from the total. Note that this is a measure of quantity because the cost of additions are being divided by the price of additions. The price is being measured by the construction cost index on line 101.
7. Review the formula for Quantity of Capital Removed. Note that it depends upon the depreciation rate and the previous year capital quantity.
8. Review the formula for Capital Quantity on line 107. Note that the formula modifies the previous year capital quantity by adding new plant and removing depreciated plant. This is the perpetual inventory method discussed elsewhere.
9. Review the Capital Cost formula on line 108. Note that it is the product of the capital price and capital quantity. Also note that the capital quantities were calculated by dividing cost by construction cost index and the capital price contains the same capital price index. Note that even though neither the price nor quantity index has recognizable

units due to the use of the construction price index, when multiplied together the units of the construction price index effectively “cancel” and the result is a cost in dollars.[9]

10. Review the total cost calculation and observe that it is the sum of OM&A and Capital Cost.

At this point “actual” total cost should be validated. The actual total cost figure calculated using the above methods should not be interpreted as being comparable to the accounting cost that the distributor calculates for reporting purposes nor should it be seen as a revenue requirement. The OM&A calculation contains adjustments for the sake of obtaining a definition of cost that allows better comparisons among distributors. The capital cost is calculated using gross plant and gross additions data from the distributor, but does not use distributor data for other capital related items such as depreciation and interest payments on debt. These items have been standardized across distributors with the goal that the market conditions under which debt was incurred and the age of plant will not influence the determination of relative cost performance. The level of cost obtained by these methods is not relevant in isolation. It is in relation to the cost levels for other distributors using the same methodology that they become meaningful.

4 Predicted Cost

4.1 Background and Theory of the Predicted Cost Calculations

The calculation of predicted cost involves the use of identified business conditions to construct variables consistent with the econometric model estimated as part of IRM-4. The business conditions required include measures of LDC output, input prices for capital and OM&A, and other variables to measure the size and age of the system. Once constructed, these variables are multiplied by the parameters of the econometric model to obtain a measure of predicted cost.

Conceptually, the econometric model takes cost data from all the distributors in Ontario and attempts to explain why they are different. Reasons why cost should be different include the fact the some distributors are bigger than others and some have to pay higher prices for labor. The econometric model then attempts to determine how important each of these factors are to the level of cost. The importance of each of factor on cost is called the parameter value. They are determined by what combination of parameters when multiplied by the factors best fit the observed cost data.

[9] For example, assume that there was one year’s worth of investment that was built for $10 million. If the construction cost index value was 2.00, the quantity would by 5 million units (10/2). Because the capital price index also contains the same construction cost index, it results in a price expressed in dollars per unit. The multiplication of these results in dollars.

Once estimates of how important each factor is on cost is determined, a prediction equation is determined through the regression results of the model. By entering data for a set of factors, a predicted cost can be generated consistent with those factors. The resulting prediction is the level of cost a typical distributor in Ontario would have if they had faced that particular set of factors. By entering the factors faced by each distributor into the equation, the resulting predictions represent the level of cost for a hypothetical distributor with the same values for the business conditions variables as the selected distributor. Other benchmarking studies may rely upon comparisons to groups of peers that are seen to be similar to the subject company. Using an econometric model effectively creates a peer group of one hypothetical company that has the exact same values for identified cost drivers as the subject company.

The econometric model will generate parameters that best fit the sample used to estimate the model. The cost predictions that come out of the model therefore create an average performance standard. Approximately half of the sample will be superior and the other half inferior based on this method. Performance is measured based upon how far a give company's performance differs from average. Other benchmarking methods such statistical frontier analysis or non-cost based methods to identify best practices will generate results in which few if any companies will meet the standard. Performance is measured by how far short a given company falls from ideal performance.

A comparison of the actual cost of the distributor to the hypothetical distributor will generate a difference in cost. The methodology will attribute differences in actual cost vs. the typical cost generated by the prediction equation to management performance. Those distributors that have cost levels lower than that for a typical distributor with the same values for business conditions will be classified as superior cost performers and those that do not will be called inferior cost performers. The discussion below will examine the econometric model in greater detail.

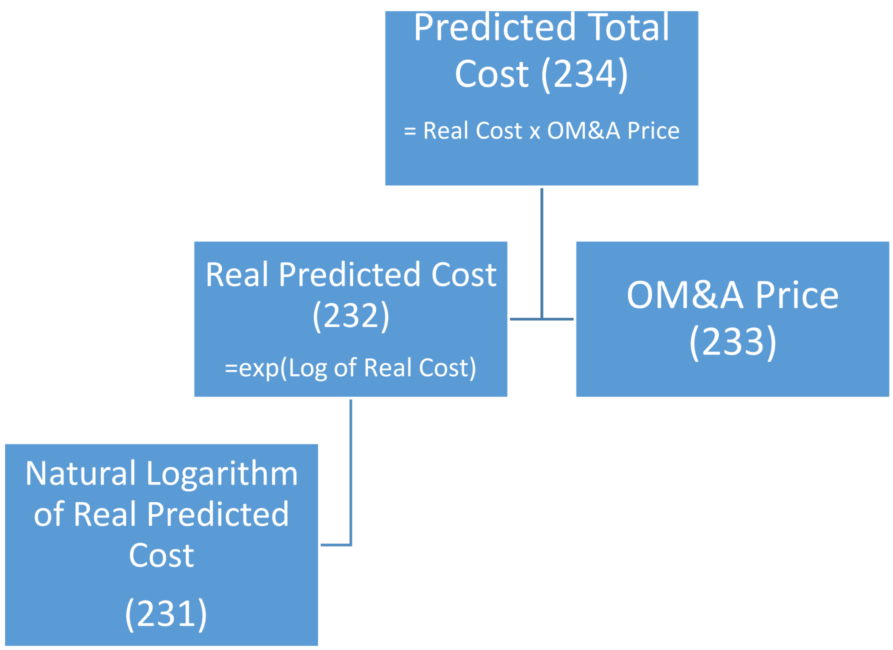
The econometric model starts with the assumption that cost (C) can be described as a linear relationship among business conditions variables (X)

$$C = B0 + B1 \times X1 + B2 \times X2 \ldots$$

The goal of the econometric work is to determine what values B best fit the observed data. The values for B are called the parameter estimates.

There are four levels of complication to consider and adjust for when adapting the data provided to work with the econometric model. The first level of complication is that the econometric model is designed to predict what economists call "real" cost which is cost adjusted for inflation[10]. This real cost prediction is then multiplied by the inflation index to obtain a prediction of normal or "nominal" cost. This procedure is standard practice in the estimate of cost functions and it guarantees a logical relationship among certain parameters of the econometric model. The second complication is that most of the variables are transformed by taking the natural logarithm. This is done because production is more of a multiplicative (i.e. XY) than linear (i.e. X+Y) in nature because it is not normally possible to use, for example, zero labor and expect to get any output. The advantage of logarithms is that it converts a multiplicative relationship to a linear one. The third complication is that the theory of cost functions calls for the use of additional variables for the squared values of the input prices and output quantities and also the product of these with each other (e.g. customer variable squared, capital price x kWh delivered). The last complication is that it is standard practice to transform the data for many of the variables to express them relative to the sample mean prior to constructing variables. The following figures work backwards from predicted cost to show how the data are related. Predicted cost is the product of the OM&A price index and real predicted cost.

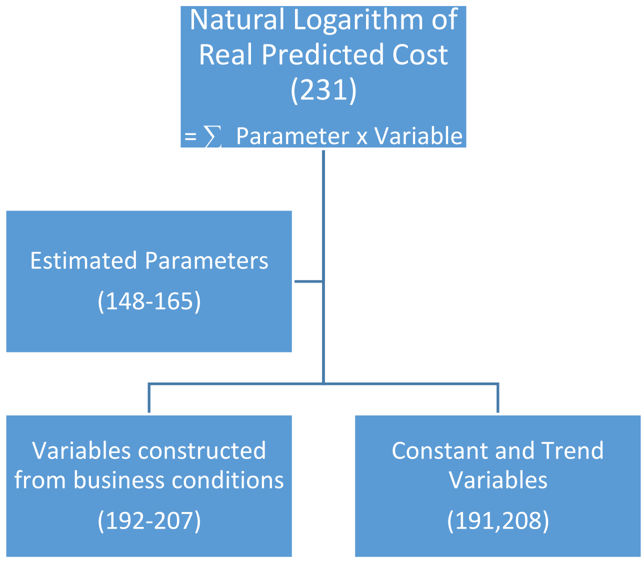
Figure 5

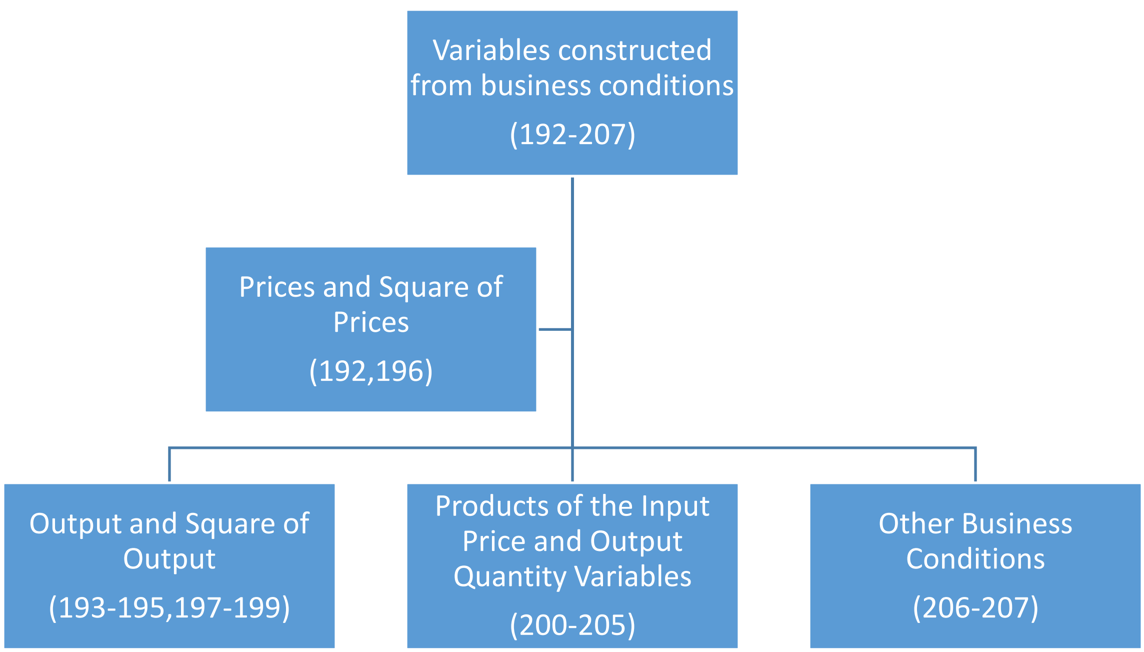


[10] For example, if wages go up by 5% and inflation is 3%, then the person has "real" wage growth of 2%. In the context of this model, dividing cost level by the OM&A price level results in a measure of real cost. This conversion is also done on the other side of the equation by dividing the capital price level by the OM&A price level which results in a "real" capital price level.

The natural logarithm of Real Predicted Cost is what is obtained when the estimated parameters are multiplied by the values of the business condition variables for an LDC. The real predicted cost is calculated by taking the exponent (anti-log) of this logarithmic value.

Figure 6





Most of the above variables are constructed by first taking the ratio of the value of an LDC to the mean for the sample used in estimating the model and then taking the natural logarithm.

The calculations on the spreadsheets are discussed below. They are done in reverse order of the previous figures. They start with the source data and eventually lead to predicted cost.

4.2 Review of the Predicted Cost Calculations on the Validation Worksheet

Section 3 of the Validation worksheet contains the calculations for predicted cost. It starts by identifying the required data on lines 114-131. The output data on lines 114 through 116 are pulled down from cells in Section 1. The capacity proxy variable is defined as the larger of the current year peak demand and all previous peak demands going back to 2002. The methodology recognizes that the capacity of the system is directly related to the highest observed peak at any point in time. The theory is that if the peak was observed, then the system had to have had enough capacity capability in place to serve customers regardless of the year in which the peak was observed. . If the peak rises to a level above the previously estimated capacity, then the system capacity must have been increased to allow the new peak to be observed. The 2012 value is gathered from the IRM-4 database and it is compared to the current year peak. The larger of the two is the capacity proxy on line 117. To validate these calculations:

1. Examine the formulas in lines 114-116 to verify they are referencing the previously validated data in Section 1.
2. Note that the 2012 value for the capacity proxy is required and taken from the IRM-4 worksheets.
3. Examine the formula on line 117 to verify that it is taking the larger of the 2012 capacity proxy and the 2013 peak demand.

The input price data are contained in lines 119 through 123. The capital price is pulled down from the calculations done in section 2. The OM&A price is calculated in a similar manner to the construction cost index in that the previous IRM-4 value is escalated. In the case of OM&A, the methodology calls for 70% of the weight to be placed on the growth of average weekly earnings and the other 30% on the Gross Domestic Product Implicit Price Deflator for Final Domestic Demand (GDPIPI FDD). The values for these price indexes are gathered from other sheets. The growth in the OM&A price index is calculated on line 122 as a weighted average of the logarithmic growth of two price indexes. The previous price index level is escalated in line 123. The values of the OM&A price index levels will differ by company. This is because local differences in price levels were incorporated into this index. These local factors were based on work done during IRM-3 and carried forward to the IRM-4 work. To validate these calculations:

1. Note that the 2012 values on lines 120, 121, and 123 are required for the calculations and these formulas reference the IRM-4 worksheets.
2. Examine the 2013 formulas on lines 120 and 121 and observe that these escalate the 2012 values using the price index data from Statistics Canada.
3. Review the formula for 2013 for the OM&A Price Index Growth on line 122. Note that this is a weighted average of the growth in the prices on lines 120 and 121.
4. Review the formula on line 123 and observe that it escalates the 2012 value by the growth rate on line 122.
5. Note that the capital price index on line 125 references the value calculated and validated in Section 2.

The other business condition variables are calculated on lines 127 through 131. The average line length and customer growth percentage are updated to include 2013 data. Previous values from the IRM-4 database are required to do this. In the case of line length, the average was extended and in the case of the customer growth it was rolled forward. This was done based on the different definitions of the variables. These variables were cross-sectional in the sense that each company had a different value that did not vary over time. To validate these calculations:

1. Note that the formula on line 128 references the data validated in Section 1.
2. Note that the 2012 value for average line km is taken from the IRM-4 worksheets and the 2013 value for line km is included in an updated average.
3. Note that the older customer data on line 130 are taken from the IRM-4 worksheets and that a new 10-year customer growth percentage is calculated from the values validated on line 113 and the older data.

At this point all the data necessary for econometric model have been calculated and validated. The variables used to construct the variables consistent with the econometric model are gathered in lines 135 through 144. The capital price is divided by the materials price to obtain the “real” price of capital. A constant and trend variable are also added. To validate these calculations:

1. Note that a constant and trend variable have been created with values that are not calculated and are therefore the same for every LDC.

2. Review the formula for the real capital price on line 138. Note that it references previously calculated and validated data.
3. Note that the formulas on lines 139 to 143 reference previous validated items in earlier rows.

The variables collected on lines 137 to 144 can be viewed as the inputs to the econometric model. Although these variables will be undergo some transformations and new variables based on these data will be created, the only additional information required to produce predicted cost will be the OM&A price index previously calculated[11] .

4.2.1 Predicted Cost Calculation Approach Using Blocks of Variables

The next steps in the approach of the calculations are done in 4 "blocks" that have identical items in each row of the block. Each block represents one part of a calculation that must be done for each variable in the prediction equation. For each variable the worksheet will calculate the product of the parameter and the transformed variable. It can be written as B x LN(X/Y) where parts of this equation represents the different blocks[12]. Put another way, each block presents the calculation of a specific part of the equation. The blocks are:

1. The parameter estimates for each variable in the model (under the heading "Company-Specific Parameter Estimates*"). This is "B" in the above equation.
2. The sample mean values for the variable in the model (under the heading "Sample Mean Values"). This is "Y" in the above equation.
3. The values for the LDC as transformed (under the heading "2013 Values Logged and Mean Scaled (where applicable")). This is "LN(X/Y)" in the above equation.
4. The resultant product (under the heading "Product of Parameter and 2013 Values") of the LDC values provided in the above third block and the above parameters in the first block. "This is B x LN(X/Y)" in the above equation.

Each of the four blocks be discussed below.

[11] The OM&A price index is not a direct input to the model but is included indirectly through the calculation on line 138. However, the OM&A price level will be required to convert the output of the econometric model back to regular nominal dollars from the "real" dollars calculated by the model.

[12] Several variables will not exactly follow this simple equation. The trend, constant, and customer growth variables do not have the logarithm taken and the "squared" variable are multiplied by 0.5.

The first block contains company-specific parameter estimates contained in lines 148 through 165. These values come from the output of the econometric model estimated during IRM-4. These are specific to each company. This was done by doing a separate estimation for each company. The only difference is that the subject company was excluded from the sample for this estimation. By doing this, the data for the selected company was not permitted to influence the parameters that would generate the standard by which their cost would be evaluated. These are gathered from the 2013 calculations worksheet, but could have been taken from a sheet specific to the parameters. These values come from the output of the econometric program code done as part of IRM-4.

The second block contains the sample mean values on lines 169-185. These will be used to transform the data on lines 138 through 143 to be consistent with the econometric model. These are also a product of the IRM-4 econometric program code.

The third block of calculations are contained on lines 191 through 208. The values in this block are obtained as follows:

1. The data collected in lines 138-142 will be transformed and placed in rows 191-195 and 206. The transformation is to first take the value for the company and divide by the sample mean value. After this is done, the natural logarithm is taken. The customer growth variable is transformed by taking the ratio to the sample mean, but the logarithm is not taken. The constant and trend variables are not transformed.
2. Additional variables based on the transformed data are created in lines 196-205. The first set of variables is the squares of the variables in lines 192 to 195 divided by two. The second set are the all the possible products of the variables in lines 196-205.

The reason why these transformations are required is that these are the same transformations that were done to the 2002-2012 data that the econometric used to obtain the parameter estimates in the first block. They need to be consistent so that the right values will be “plugged into” the formula and a sensible output will be obtained.

The fourth block of calculations are contained in lines 212-229. These are the line by line products of the parameters in the first block and the company values in the third block. Because the parameter value represents how important each variable is in the determination of cost and the LDC value for the variable is how different the LDC is from average, the product of the parameter and the company value can be seen as the cost impact of being different from average. The resulting sum can be seen as a weighted average of the cost impact of all the identified ways in which the LDC is different than average. Because of the manner in which transformation done the result is not yet expressed in dollars. The result is the natural logarithm of real predicted cost at this stage of the calculation. The calculations to obtain predicted cost are contained in lines 231-234. This is done by reversing the logarithm and multiplying the result by the OM&A price index that was used to obtain "real" cost.

To validate the calculations in the four blocks contained in lines 148 through 229, the following steps are suggested:

1. Note that the parameter values in the first block of rows 148 – 165 are pulled from the 2013 Benchmarking Calculations worksheet. Confirm these values are for the selected LDC. Note that these are pasted values.
2. Note that the sample average values in second block of rows (114-131) are also pasted values. Note that the source of both of the values for the first two blocks are from the previous IRM-4 work.
3. Review the formulas in the third block of rows (191-208). Observe that these follow the transformation procedure previously described.
4. Review the formulas in the fourth block of rows (212-229). Observe that in each case the value is the product of the corresponding rows in block 1 and block 2.
5. Review the formula on line 231. Verify that it sums the values in block 3.
6. Review the formula on line 232 that uses the EXP function to reverse the logarithm. If unfamiliar with logarithms, the user may wish to perform a simple calculation in column M to demonstrate how these functions work. Enter the value 5 in cell M240. Enter the formula "=LN(M240)" into cell M241 and obtain the value of 1.609438. Enter the formula "=EXP(M241)" into cell M242 and obtain the value of 5. In other words EXP(LN (X)) = X and therefore the EXP function reverses the LN function.
7. Note that the previously calculated and validated OM&A price is "pulled down" from above via formula on line 233.
8. Review the formula on line 234 for Predicted Cost. Note that by multiplying the result on line 232 by the OM&A price the formula reverses the effect of the calculation done on line 138 in which the capital price was divided by the OM&A price. The calculation on line 138 effectively made the model predict "real" cost as described above. The calculation on line 234 effectively converts the "real" cost back to the familiar "nominal" cost.

Section 4 of the Validation worksheet contains the cost performance results. The cost performance of the LDC is the logarithmic difference between actual and predicted cost. This is done in the last section of the worksheet on lines 237-250. Because calculating percentage differences using logarithms in unfamiliar to most people, the familiar arithmetic version is also included for comparison. The logarithmic method is used because the arithmetic is known to produce biased results when averaging results over multiple periods.

Published results for previous years are contained on lines 248 and 249. These will allow the calculation of three year average performance. This performance is the measure used to determine the stretch factor assignment. To validate the formulas in Section 4:

1. Verify that the formulas in lines 237 and 238 reference the previously calculated Actual and Predicted Cost.
2. Examine the formula in line 242. This formula calculates the percentage difference logarithmically. Compare this to the formula on line 240 that does the more familiar arithmetic calculation.
3. Verify the formula on line 250 that calculates the new three year average performance using results from 2011-2013.

5 Forecasting Future Cost Performance

The worksheet titled Forecasting contains the formulas necessary to forecast future benchmarking results. The worksheet will look very similar to the Validation sheet except the formulas are stretched forward to 2014 and beyond. To obtain results, forecasts of cost and the necessary business conditions need to be generated. The data required are:

- OM&A cost
- Gross plant additions (net of HV)
- Number of Customers
- Delivery Volume
- Peak Demand
- Km of Line
- Labor price inflation
- Economy wide inflation
- Construction cost inflation
- Rate of return allowed by the OEB

Default values for the required data items were generated from assumed rates of growth. These assumptions are contained on the worksheet titled Assumptions for Forecasting. The values in these cells are arbitrary and should be changed by the user.

These assumed rates of growth are used to generate future data. By changing the assumed growth rates, different values consistent with the new values will be generated. This method was used because it was the easiest method to generate future data that would work for any selected LDC. These growth rate assumptions may be improved in several ways. These include:

- obtaining actual values or better forecasts for input price variables
- The LDC entering a better estimated growth of OM&A and gross plant additions
- The LDC entering the account level OM&A values and gross plant additions in place of those generated by escalating historical values

It is recommended that once the user is comfortable with the way the model works, the formulas to generate the future data should be replaced with the distributor's own accounting data where applicable.

The escalation method is reasonable to obtain 2014 values, but it is not likely that this method would produce accurate values for each year of a multi-year forecast period. It is therefore recommended that anyone wishing to produce forecasts beyond 2014 enter the values for each year separately based on their own forecast models. The rows that require the user to provided data are highlighted in grey. These currently contain formulas to generate default values, but may be replaced.

6 Overview of Spreadsheets

In order to give the user a better understanding of how all the worksheets in the 2013 working papers relate to the calculations, a flowchart is included on the worksheet titled "Overview of Worksheets". This flowchart groups the various worksheet into categories based upon the source of the data. The categories are as follows:

- Previous IRM-4 work by PEG
- Statistics Canada
- Provided by LDCs
- Supplemental Data Provided by LDCs

Each will be discussed in turn.

It was necessary to draw upon previous IRM-4 work in several cases. The values used for the price indexes must be consistent with the previous work in order to provide correct inputs to the model. These prices include those for average hourly earnings, the construction cost index, and the GDP-IPI. The previous values are escalated by the most recent information. The capital data are also necessary due to the nature of the perpetual inventory method discussed above. One of the worksheets (2012 BM Database) is an extract of the larger database to allow for easier calculations. The other two worksheets are taken from the IRM-4 working papers.

The three price indexes from Statistic Canada are identified and each has its own associated worksheet. The remaining worksheets contain LDC data. With the exception of the HON LV charges data, these data were compiled by OEB staff based on data provided by LDCs. The HON LV data required some processing in order to be compatible with the format of the benchmarking calculations. These are contained on the LV Pivot and LV worksheets. Some formatting was also required for the "Acct 5014 5015 5112" data which is done in the HV Charges spreadsheet.

7 The Generic LDC Worksheet

A worksheet titled "Generic LDC Worksheet" is included to help illustrate how the predicted cost calculations work. It is intended to be a "whiteboard" that can be used to observe how predicted cost changes as a result of changing the values of business conditions. It has been put on a separate sheet so that the user may experiment with the model without altering any of the calculations done elsewhere. It is "Generic" in the sense that it is not specific to any particular LDC. By default, it is populated with average business condition values for all benchmarked Ontario LDCs and parameters values based on the full sample of LDCs. As such, this worksheet is limited in application in that it is only representative of a hypothetically average LDC and not that of a specific distributor.

One feature of the data transformations described in section 4.1 is that the calculations become much less complex in the special case in which sample average values are used. This is because these ratios for the mean company are 1.00 and this becomes zero when the natural log is taken. This will effectively make most of the terms of the prediction equation equal to zero. The remaining terms are the constant, the customer growth variable, and the trend variable. This property is the main reason why the data are transformed prior to model estimation.

Overall the Generic LDC Worksheet will look very similar to the Validation worksheet starting at Section 3. The values used are taken from a calculation of the industry average for each required variable. The effect of using these average values can be seen on lines 192 to 206 of the worksheet. Note that the values for many of the variables have become zero. This effectively removes them from the calculation in this case and the prediction equation is now only dependent upon the constant, average line length, and trend variables on lines 191, 207, and 208.

By experimenting with the values of the business conditions, one can see the impact on predicted cost. This can be used as a tool to better understand how the model works. It could also be a tool to allow a step-by-step modification of the data to trace the cost prediction impact of each business condition as it becomes different from “average”. A column with the sample mean values is provided to restore these values as well as formulas to escalate the output variables. The bottom of the calculations contains the value of the cost prediction from using mean values so that the impact of changes can be seen. This worksheet is supplementary to the other work. It is designed to allow experimentation with the model without modifying the work done to obtain results. It is optional and can be omitted from the review process if desired.

As an illustration, enter 10% into cell AP129. The value used for the number of customers will change to 70,092 which is 10% higher than the average value[13]. For example, if the number of customers is arbitrarily increased to be 10% higher than the default sample average provided, one can see that the values in lines 193 and 197 are no longer zero. These values will now contribute to the calculation of predicted cost because customers are no longer average. The result on line 238 indicates the predicted cost is now 45,398,636 instead of 43,507,203 which is 4.26% higher than before the customer value was increased.

8 Error Correction and Future Improvements

This spreadsheet model was designed to help LDC staff obtain a better understanding of the calculations done to obtain the 2013 benchmarking results. It contains corrections by PEG and approved data corrections from a few LDCs. The results will therefore not match those contained in the 2014 PEG report. Any additional data revisions or other issues regarding the calculations will be handled as part of the 2014 data update later this year.

[13] These percentage differences are calculated using logarithms which will be similar to the more familiar arithmetic calculation of ((A/B) – 1).

Possible future improvements to this work include consolidating the IRM-4 working papers into the spreadsheet or moving the calculation backwards in time. Other possible improvements include work to obtain a more detailed analysis of the cost performance results.

It is a major objective of this work to make these complex calculations as accessible and transparent as possible. It is hoped that this work will answer LDC questions and raise confidence in the benchmarking process.

Appendix 1: Key Documents

Benchmarking Documents Associated with IRM-4

Report of the Board – A Renewed Regulatory Framework for Electricity Distributors: A Performance Based Approach (Oct 18, 2012)
http://www.ontarioenergyboard.ca/oeb/_Documents/Documents/Report_Renewed_Regulatory_Framework_RRFE_20121018.pdf

"Empirical Research in Support of Incentive Rate-Setting: 2013 Benchmarking Update".

http://www.ontarioenergyboard.ca/oeb/_Documents/EB-2010-0379/PEG_Benchmarking_Report_20140814.pdf

Memorandum to Ontario Energy Board Staff

http://www.ontarioenergyboard.ca/oeb/_Documents/EB-2010-0379/EB-2010-0379_PEG_Memorandum_Corrections_20131219.pdf

Productivity and Benchmarking Research in Support of Incentive Rate Setting in Ontario: Final Report to the Ontario Energy Board

http://www.ontarioenergyboard.ca/oeb/_Documents/EB-2010-0379/EB-2010-0379_Final_PEG_Report_20131111.pdf

PEG IRM-4 Working Papers Documentation and Files

http://www.ontarioenergyboard.ca/oeb/_Documents/EB-2010-0379/EB-2010-0379_Documentation_PEG_Working_Papers.pdf

http://www.ontarioenergyboard.ca/oeb/_Documents/EB-2010-0379/EB-2010-0379%20PEG%20TFP%20and%20BM%20database%20calculations.xlsx

Benchmarking Documents Prior to IRM-4

Benchmarking the Costs of Ontario Power Distributors

(http://www.ontarioenergyboard.ca/documents/cases/EB-2006-0268/PEG_Final_Benchmarking_Report_20080320.pdf

Appendix 2: Additional Background on the Indexing Logic

The Price Cap IR will use a price cap index (PCI) formula to restrict the change in electricity distribution prices. While PCIs vary from plan to plan, the PCI growth rate $(growthPCI)$ is typically given by the growth in an inflation factor (*P*) minus an X-factor (*X*) plus or minus a Z-factor (*Z*), as in the formula below:

$$growth\ PCI = P - X \pm Z.$$

In this formula, the P is the inflation in the prices of inputs, the X is the productivity growth expected of the LDC and Z is a possible adjustment for unforeseen event such as natural disasters.

The logic of the P – X part of the formula can be derived from a decomposition of distributor cost. Cost (C) is equal to the product of the quantity of inputs (X) used and the price of those inputs (W). It is also true that the cost per unit output (Y) is equal to inputs times input price divided by output. Because total factor productivity (TFP) is equal to output quantity per unit of input quantity (Y / X), unit cost is equal to input price divided by productivity.

C = W x X (by definition)

C/Y = W x X / Y

TFP = Y / X (by definition)

C/Y = W / TFP

The growth of unit cost is equal to the growth in input prices less the growth in TFP. This is the basis for the IPI – X formula which contains the growth in input prices (P) and expected productivity growth (X). The X in this formula is expected productivity which will be estimated as the sum of a long run TFP trend and a stretch factor to capture expected productivity in excess of the trend. The reason why this is an appropriate formula to adjust rates is that rates are equal to revenue generated per unit of output. It is appropriate that this should equal the cost per unit of producing that output. A more formal presentation of this logic is presented below.

The following material has been taken from the section 2 of the PEG IRM-4 report titled "PRODUCTIVITY AND BENCHMARKING RESEARCH IN SUPPORT OF INCENTIVE RATE SETTING IN ONTARIO: FINAL REPORT TO THE ONTARIO ENERGY BOARD".

Indexing Logic

The Price Cap IR will use a price cap index (PCI) formula to restrict the change in electricity distribution prices. While PCIs vary from plan to plan, the PCI growth rate $(growthPCI)$ is typically given by the growth in an inflation factor (*P*) minus an X-factor (*X*) plus or minus a Z-factor (*Z*), as in the formula below:

$$growth\,PCI = P - X \pm Z. \quad [1]$$

In North American regulation, the terms of the PCI are set so that the change in regulated prices mimics how prices change, in the long run, in competitive markets. This is a reasonable basis for calibrating utility prices since rate regulation is often viewed as a surrogate for the competitive pressures that would otherwise lead to "just and reasonable" rates. Economic theory has also established that competitive markets often create the maximum amount of benefits for society.[14] It follows that effective utility regulation should replicate, to the greatest extent possible, the operation and outcomes of competitive markets. A "competitive market paradigm" is therefore useful for establishing effective regulatory arrangements, and several features of competitive markets have implications for how to calibrate PCI formulas.

One important aspect of competitive markets is that prices are "external" to the costs or returns of any individual firm. By definition, firms in competitive markets are not able to affect the market price through their own actions. Rather, in the long run, the prices facing any competitive market firm will change at the same rate as the growth in the industry's unit cost.

Competitive market prices also depend on the *average* performance in the industry. Competitive markets are continually in a state of flux, with some firms earning more and others less than the "normal" rate of return on invested capital. Over time, the average performance exhibited in the industry is reflected in the market price.[15]

---

[14] This is sometimes known as the "First Fundamental Welfare Theorem" of economics, but it should be noted that the theoretical finding that competition leads to efficient outcomes does not apply under all conditions (*e.g.* if there are externalities whose costs or benefits are not reflected in competitive market prices).

[15] This point has also been made in the seminal 1986 article in the Yale Journal of Regulation, *Incentive Regulation for Electric Utilities* by P. Joskow and R. Schmalensee. They write "at any instant, some firms (in competitive markets) will earn more a competitive return, and others will earn less. An efficient competitive firm

Taken together, these features have the important implication that in competitive markets, returns are commensurate with performance. A firm can improve its returns relative to its rivals by becoming more efficient than those firms. Companies are not disincented from improving efficiency by the prospect that such actions will be translated into lower prices because the prices facing any individual firm are external to its performance. Firms that attain average performance levels, as reflected in industry prices, would earn a normal return on their invested capital. Firms that are superior performers earn above average returns, while firms with inferior performance earn below average returns. Regulation that is designed to mimic the operation and outcomes of competitive markets should allow for this important result.

Another implication of the competitive market paradigm bears a direct relationship to the calibration of PCI formulas. As noted above, in the long run, competitive market prices grow at the same rate as the industry trend in unit cost. Industry unit cost trends can be decomposed into the trend in the industry's input prices minus the trend in industry total factor productivity (TFP). Thus if the selected inflation measure is approximately equal to the growth in the industry's input prices, the first step in implementing the competitive market paradigm is to calibrate the X factor using the industry's long-run TFP trend.

The mathematical logic underlying this result merits explanation. We begin by noting that if an industry earns a competitive rate of return in the long run, the growth in an index of the prices it charges (its output prices) will equal its growth in unit cost.

$$trend\ Output\ Prices^{Industry} = trend\ Unit\ Cost^{Industry}. \quad [2]$$

As stated above, the trend in an industry's unit cost is the difference between trends in its input price index and its TFP index. The full logic behind this result is presented below:

$$\begin{aligned}
\text{trend Unit Cost}^{\text{Industry}} &= \text{trend Cost}^{\text{Industry}} - \text{trendOutput Quantities}^{\text{Industry}} \\
&= \left(\textit{trend}\ \text{Input Prices}^{\text{Industry}} + \textit{trend}\ \text{Input Quantities}^{\text{Industry}}\right) \\
&\quad - \textit{trend}\ \text{Output Quantities}^{\text{Industry}} \\
&= \textit{trend}\ \text{Input Prices}^{\text{Industry}} \\
&\quad - \left(\text{trend Output Quantities}^{\text{Industry}} - \textit{trend}\ \text{Input Quantities}^{\text{Industry}}\right) \\
&= \text{trend Input Prices}^{\text{Industry}} - \textit{trend}\ \text{TFP}^{\text{Industry}}.
\end{aligned} \quad [3]$$

will expect on average to earn a normal return on its investments when they are made, and in the long run the average firm will earn a competitive rate of return"; *op cit*, p. 11.

Substituting [3] into [2] we obtain

$$trend\ Output\ Prices^{Industry} = trend\ Input\ Prices^{Industry} - trend\ TFP^{Industry} \qquad [4]$$

Equation [4] demonstrates the relationship between the X factor and the industry TFP trend. If the selected inflation measure (*P* in equation [1]) is a good proxy for the industry's trend in input prices, then choosing an X factor equal to the industry's TFP trend causes output prices to grow at the rate that would be expected in a competitive industry in the long run. This is the fundamental rationale for using information on TFP trends to calibrate the X factor in index-based PBR plans.

It should be emphasized that both the input price and TFP indexes above correspond to those for the relevant utility *industry*. This is necessary for the allowed change in prices to conform with the competitive market paradigm. In competitive markets, prices change at the same rate as the industry's trend in unit costs and are not sensitive to the unit cost trend of any individual firm. This is equivalent to saying that competitive market prices are external to the performance of any given firm in the industry.

There are two main options for selecting inflation factors in index-based PBR plans. One general approach is to use a measure of economy-wide inflation such as those prepared by government agencies. Examples include the Gross Domestic Product Implicit Price Index (GDP-IPI) or the US Price Index for Gross Domestic Product (GDP-PI). An established alternative is to construct an index of external price trends for the inputs used to provide utility services. This approach is explicitly designed to measure input price inflation of the regulated industry.[16] The Board has found that the inflation factor in Price Cap IR will be closer to a measure of industry input price inflation, so the indexing logic presented in equations [1] through [4] is valid for Price Cap IR.

While industry TFP and input price measures are used to calibrate a PCI, in most index-based incentive regulation plans the X factor is greater than what is reflected in the utility industry's long-run TFP trend. This is because industry TFP trends are usually measured using historical data from utility companies. Utilities have historically not operated under the competitive market pressures that naturally create incentives to operate efficiently, and it is also widely believed that traditional, cost of service regulation does not promote efficient utility behavior.

Incentive regulation is designed to strengthen performance incentives, which should in turn encourage utilities to increase their efficiency and register more rapid TFP growth relative to historical norms. It is

[16] A less common approach is to set inflation measures using changes in *output* prices charged by peer utilities. It is important for any such peer-price inflation measure to be constructed carefully so that it reflects the circumstances of companies that are very similar to the utility subject to the incentive regulation plan.

also reasonable for these performance gains to be shared with customers since incentive rate-setting is designed to produce "win-win" outcomes for customers and shareholders. For this reason, nearly all North American incentive regulation plans have also included what are called "consumer dividends" or productivity "stretch factors" as a component of the X factor. The stretch factor reflects the expected acceleration in TFP relative to historical TFP trends.[17]

[17] More precisely, the stretch factor is that portion of the expected acceleration of TFP growth that it passed through to the change in customer rates as a form of benefit-sharing under the plan.

Appendix 3: Additional Background on the Econometric Cost Model

The following material has been taken from the section 6 and appendix 2 of the PEG IRM-4 report titled "PRODUCTIVITY AND BENCHMARKING RESEARCH IN SUPPORT OF INCENTIVE RATE SETTING IN ONTARIO: FINAL REPORT TO THE ONTARIO ENERGY BOARD".

Econometric Research on Cost Performance

Total Cost Econometric Model

PEG benchmarked the total cost of Ontario's electricity distributors using a total cost econometric model. An econometric cost function is a mathematical relationship between the cost of service and business conditions. Business conditions are aspects of a company's operating environment that may influence its costs but are largely beyond management control. Economic theory can guide the selection of business condition variables in cost function models.

According to theory, the total cost of an enterprise depends on the amount of work it performs - the scale of its output - and the prices it pays for capital goods, labor services, and other inputs to its production process.[18] Theory also provides some guidance regarding the nature of the relationship between outputs, input prices, and cost. For example, cost is likely to rise if there is inflation in input prices or more work is performed.

For electricity distribution, total customers served and total kWh delivered are commonly used for output variables. Peak demand is another potential output variable. Peak demand is a billing determinant for some customers, but peak demand will also be an important cost driver for smaller customers whose peak demands are not metered. The reason is that delivery systems must be sized to accommodate peak demands, so there is a direct relationship between customers' peak demands and the costs of the necessary power delivery infrastructure.

In addition to output quantities and input prices, electricity distributors confront other operating conditions due to their special circumstances. Unlike firms in competitive industries, electricity distributors are obligated to provide service to customers within a given service territory. Distribution services are delivered directly into the homes, offices and businesses of

[18] Labor prices are usually determined in local markets, while prices for capital goods and materials are often determined in national or even international markets.

end-users in this territory. Distributor cost is therefore sensitive to the circumstances of the territories in which they provide delivery service.

One important factor affecting cost is customer location. This follows from the fact that distribution services are delivered over networks that are linked directly to customers. The location of customers throughout the territory directly affects the assets that utilities must put in place to provide service. The spatial distribution of customers will therefore have implications for network cost.

The spatial distribution of customers is sometimes proxied by the total circuit km of distribution line, or the total square km of territory served. Provided customer numbers is also used as a cost measure, these variables will together reflect the impact of different levels of customer density within a territory on electricity distribution costs.

Cost can also be sensitive to the mix of customers served. The assets needed to provide delivery service will differ somewhat for residential, commercial, and industrial customers. Different types of customers also have different levels and temporal patterns of demand and different load factors.

In addition to customer characteristics, cost can be sensitive to the physical environment of the service territory. The cost of constructing, operating and maintaining a network will depend on the terrain over which the network extends. These costs will also be influenced by weather and related factors. For example, costs will likely be higher in areas with a propensity for ice storms or other severe weather that can damage equipment and disrupt service. Operating costs will also be influenced by the type and density of vegetation in the territory, which will be at least partly correlated with precipitation and other weather variables.

Econometric cost functions require that a functional form be specified that relates cost to outputs, input prices, and other business conditions. The parameter associated with a given variable reflects its impact on the dependent cost variable. Econometric methods are used to estimate the parameters of cost function models. Econometric estimates of cost function parameters are obtained using historical data on the costs incurred by distributors and measurable business condition variables that are included in the cost model.

Econometric Research on Electricity Distribution Cost

Economic theory says that the cost of an enterprise depends on input prices and the scale of output. PEG's cost function included input prices, as defined and measured in Chapter Three of this report. PEG investigated a number of different choices for output variables, including customer numbers, kWh deliveries, different measures of peak demand, and total km of line. We also investigated the impact of other business condition variables that are largely beyond management control but can still impact distribution cost. Data on both the output and business condition variables were drawn from Section 2.1.5 of the RRRs.

PEG consulted with stakeholders extensively on the choices for outputs and business condition variables in the econometric work. This included discussions with the PBR Working Group, as well as a March 1, 2013 webinar on the topic in which the entire industry and other stakeholders were invited to participate. These consultations examined the merits of a variety of "cost driver" variables that PEG considered during its econometric work. In addition to outputs, the business condition variables explored could be categorized as belonging to one of five sets of cost drivers:

1) The mix of customers served *e.g.* serving a more industrialized customer base, load factor;
2) Variables correlated with urbanization and urban density, such as municipal population per square km of urban territory, the percent of urban territory in total territory, or the share of lines that are underground ;
3) Geography, such as total area served, the share of territory that is on the Canadian shield, and whether a distributor's territory is in Northern Ontario;
4) The age of assets, as proxied by accumulated depreciation relative to gross plant value or the share of total customers that were added in the last 10 years; or
5) High-voltage intensiveness, such as the share of transmission substation assets (greater than 50 kV) in total distribution plant. This variable was designed to reflect costs associated with high voltage assets that could not be specifically identified and eliminated from our cost measure.

The total cost benchmarking model also contains a trend variable. This variable captures systematic changes in costs over time that are not explained by the specified business conditions. It may also reflect the failure of the included business condition variables to measure the trends in relevant cost drivers properly. The model may, for instance, exclude an important cost driver or measure such a cost driver imperfectly. The trend variable might then capture the impact on cost of the trend in the driver variable.

Estimation Results and Econometric Benchmarking

Econometric Results

Estimation results for our electricity distribution cost model are reported in Table 16. The estimated coefficients for the business conditions and the "first order" terms of the output variables are elasticities of cost for the sample mean firm with respect to the variable. The first order terms do not involve squared values of business condition variables or interactions between different variables.

Table 16 also reports the t-statistic values generated by the estimation program. The t-statistic values were used to assess the statistical significance of the estimated cost function parameters. A parameter estimate is deemed statistically significant if the hypothesis that the true parameter value equals zero is rejected at a 5% significance level (*i.e.* a 95% confidence level). Each statistically significant parameter estimate is identified with an asterisk.

Examining the results in Table 16, it can be seen that there are three statistically significant output variables: customer numbers; kWh deliveries; and system capacity peak demand. Our measure of customer numbers is equal to total customers minus street lighting, sentinel lighting, and unmetered scattered load. The kWh deliveries measure is billed kWh deliveries (before loss adjustment) to all customers.

The system capacity peak demand measure was equal to the highest annual peak demand measure for a distributor up to the year in question. For example, in 2002 (the first sample year), the system capacity measure for each distributor was its annual peak demand for 2002. In 2003, if the distributor's reported annual peak exceeded its 2002 peak, the system capacity peak was equal to the annual peak demand in 2003. If the annual peak in 2003 was below the annual peak in 2002, the annual peak in 2002 was the highest peak demand measure reported by the distributor, and this value is therefore also recorded as the system capacity peak for 2003. Values in subsequent years were calculated in the same manner. The system capacity variable is intended to reflect distribution infrastructure sized to meet peak demands. Even if those demands fall over time, the distributor's infrastructure and its associated costs will (in nearly all cases) remain. The system capacity peak variable was suggested in the PBR Working Group discussions and largely supported by the Group.

The output parameter estimates, as well as the parameter estimate for capital input prices, were plausible as to sign and magnitude. Cost was found to increase for higher values of capital service prices and output quantities. At the sample mean, a 1% increase in the number of customers raised cost by .44%. A 1% increase in kWh deliveries raised cost by about .10%. A 1% increase in system capacity increased distribution cost by 0.16%. Customer numbers is therefore the dominant output-related cost driver, followed by peak demand, followed by kWh deliveries.

Two other business condition variables are also identified as statistically significant cost drivers on Table 16: average circuit km of line; and share of customers added over the last 10 years.

With respect to a distributor's average circuit km of line over the 2002-2012 period, it can be seen that a 1% increase in average circuit km raised distribution cost by 0.29%. PEG used average km over the sample period, rather than each distributor's reported time series of km, because of anomalous trends in circuit km data for some distributors. The circuit km coefficient therefore reflects the cost impact of cross-sectional differences in circuit km across distributors, but not the impact of *changes* in km of line (all else equal) over the 2002-2012 period, on distribution cost.

The circuit km variable clearly has an output-related dimension, because it reflects customers' location in space and distributors' concomitant need to construct delivery systems that transport electrons directly to the premises of end-users. The average circuit km variable can be considered a legitimate output when examining cross-sectional differences in costs across Ontario distributors. Circuit km could, for example, play an important role in identifying appropriate peer groups for unit cost comparisons, since this benchmarking exercise compares unit costs across Ontario distributors at a given point in time. However, it would not be appropriate for the average circuit km variable to be used as an output variable in the current TFP study. This study is designed to estimate *trends* in TFP for the Ontario electricity distribution industry, but the current average km variable only reflects cross sectional, and not trend, impacts on distribution cost.

With respect to the share of a distributor's customers that was added over the last 10 years, the variable is designed to proxy recent growth and the age of distribution systems. All else equal, serving a relatively fast-growing territory requires a greater amount of more current capital additions. These investment pressures could put upward pressure on costs. Our model shows that a 1% increase in this variable increases distribution costs by 0.017%.

A surprising finding of our cost model was the coefficient on the trend variable. This coefficient was estimated to be 0.017%. This implies that, even when input prices, outputs, and other business condition variables remain unchanged, costs for the Ontario electricity distribution industry still increased by an average of 1.7% per annum between 2002 and 2012. This is counter to the usual finding in cost research, where the coefficient on the trend variable is negative. One factor that could be contributing to these upward cost pressures is government policy implemented over the sample period. Another possibility is that there are cost pressures for a sizeable portion of the industry due to company-specific factors, rather than industry-wide policies, but it is difficult to capture these company-specific cost pressures in measurable business condition variables.

PEG did examine a wide range of other business condition variables in our cost research. These other variables were either not statistically significant or did not have sensible signs. These variables included:

- The percent of distribution territory on the Canadian shield;
- A dummy variable for whether or not a distributor was located in Northern Ontario;
- The share of transmission substation plant (greater than 50 kV) in total gross plant;
- The share of deliveries to residential customers;
- Load factor;
- The share of service territory that is urban;
- Municipal population divided by km^2 of urban territory; and
- The percentage of circuit km that are underground.

Econometric Benchmarking

PEG used its recommended cost model presented in Table 16 to generate econometric evaluations of the cost performance of Ontario electricity distributors. This was done by inserting values for each distributor's output and business condition variables into a cost model that is “fitted” with the coefficients presented in Table 16. This process yields a value for the predicted (or expected) costs for each distributor in the sample given the exact business condition variables faced by that distributor.

Econometric Research

A.2.1 Form of the Cost Model

The functional form selected for this study was the translog.[19] This very flexible function is the most frequently used in econometric cost research, and by some account the most reliable of several available alternatives.[20] The general form of the translog cost function is:

$$\begin{aligned} \ln C = \alpha_0 + \sum_h \alpha_h \ln Y_h + \sum_j \alpha_j \ln W_j \\ + \frac{1}{2}\left(\sum_h \sum_k \gamma_{h,k} \ln Y_h \ln Y_k + \sum_j \sum_n \gamma_{j,n} \ln W_j \ln W_n \right) \\ + \sum_h \sum_j \gamma_{h,j} \ln Y_h \ln W_j \end{aligned} \qquad [A2.1]$$

where Y_h denote one of K variables that quantify output and the W_j denotes one of N input prices.

One aspect of the flexibility of this function is its ability to allow the elasticity of cost with respect to each business condition variable to vary with the value of that variable. The elasticity of cost with respect to an output quantity, for instance, may be greater at smaller values of the variable than at larger values. This type of relationship between cost and quantity is often found in cost research.

Business conditions other than input prices and output quantities can contribute to differences in the costs of LDCs. To help control for other business conditions the logged values of some additional explanatory variables were added to the model in Equation [A2.1] above.

The econometric model of cost we wish to estimate can then be written as:

---

[19] The transcendental logarithmic (or translog) cost function can be derived mathematically as a second order Taylor series expansion of the logarithmic value of an arbitrary cost function around a vector of input prices and output quantities.

[20] See Guilkey (1983), et. al.

$$\ln C = \alpha_o + \sum_h \alpha_h \ln Y_h + \sum_j \alpha_j \ln W_j$$
$$+ \quad \frac{1}{2}\left[\sum_h \sum_k \gamma_{hk} \ln Y_h \ln Y_k + \sum_j \sum_n \gamma_{jn} \ln W_j \ln W_n\right] \qquad \text{[A2.2]}$$
$$+ \quad \sum_h \sum_j \gamma_{hj} \ln Y_h \ln W_j + \sum_h \alpha_h \ln Z_h + \alpha_t T + \varepsilon$$

Here the Z_h's denote the additional business conditions, T is a trend variable, and ε denotes the error term of the regression.

Cost theory requires a well-behaved cost function to be homogeneous in input prices. This implies the following three sets of restrictions:

$$\sum_{h=1}^{N} \frac{\partial \ln C}{\partial \ln W_h} = 1 \qquad \text{[A2.3]}$$

$$\sum_{h=1}^{N} \frac{\partial^2 \ln C}{\partial \ln W_h \partial \ln W_j} = 0 \qquad \forall j = 1,\ldots, N \qquad \text{[A2.4]}$$

$$\sum_{h=1}^{N} \frac{\partial^2 \ln C}{\partial \ln Y_h \partial \ln Y_j} = 0 \qquad \forall j = 1,\ldots, K \qquad \text{[A2.5]}$$

Imposing the above $(1 + N + K)$ restrictions implied above allow us to reduce the number of parameters that need be estimated by the same amount. Estimation of the parameters is now possible but this approach does not utilize all information available in helping to explain the factors that determine cost. More efficient estimates can be obtained by augmenting the cost equation with the set of cost share equations implied by Shepard's Lemma. The general form of a cost share equation for a representative input price category, *j*, can be written as:

$$S_j = \alpha_j + \sum_h \gamma_{h,j} \ln Y_h + \sum_n \gamma_{jn} \ln W_n \qquad \text{[A2.6]}$$

We note that the parameters in this equation also appear in the cost model. Since the share equations for each input price are derived from the first derivative of the translog cost function with respect to that input price, this should come as no surprise. Furthermore, because of these cross-equation restrictions, the total number of coefficients in this system of equations will be no larger than the number of coefficients required to be estimated in the cost equation itself.

Appendix 4: Glossary of Terms

Actual Cost: Actual Cost equals the sum of OM&A accounts selected for benchmarking, adjustments to OM&A expenses, and a measure of capital cost using the perpetual inventory method and a capital price index that includes standardized depreciation rates and rate of return.

Benchmarking: Benchmarking, Cost Benchmarking, and Statistical Cost Benchmarking are used interchangeably. They represent the process by which management performance in inferred by comparing a measure of actual cost to that predicted by an econometric model.

Business Condition: The term business condition denotes a measurable condition that can plausibly have an impact on the cost of power distribution. Output quantities and input prices are business conditions that have their own names due to their importance. Other business conditions identified by the model are the average km of line and 10-year customer growth. Some business conditions such as the area of the service territory served or the percent of line that is buried underground were tested and found to not have a statistically significant relationship to cost.

Econometric Cost Model: An econometric cost model (model, econometric model, statistical cost model) is a linear equation or series of equations that describe a hypothetical relationship between cost and a number of business conditions. It has a certain functional form in which cost (or a function of cost) proposed to be equal to a series of parameters multiplied by business condition variables (or functions of business condition variables). An example of "functions of" are the data transformations detailed in the body of this report. Regression analysis is used to obtain estimates for the unknown parameters. These estimated parameter values for the proposed business conditions are evaluated for statistical significance. A good model will have statistically significant estimates for the parameter values associated with each business condition.

Performance: Performance or Cost Performance is the logarithmically calculated difference between an "actual cost" specified for benchmarking purposes and that predicted by the econometric model. Because this is calculated as a residual, it will also implicitly include the impact of any business condition not reflected by other variables in the model.

Parameter: A parameter is a part of an equation with a constant value that is not known. In the equation: Cost = A + B x Customers, both cost and customers are known. What is not known are the values of the parameters A and B. Once values for A and B are estimated using econometrics, they will establish the relationship between customers and cost.

Parameter Value: The term parameter value or estimated parameter is used to indicate the specific value estimated for the unknown parameter.

Statistically Significant: Statistically significant means that uncertainty associated with the value the model estimated for a parameter is low. More formally, it means that given the variance of the underlying data, the probability that an estimated parameter is actually zero (i.e. no relationship to cost) is under 10%. This only refers to the confidence the model has that the estimate parameter is not zero. It does not imply that the level of the parameter implies a strong relationship between this business condition and cost. For example, assume that the number of Tim Horton's served by a distributor put in the econometric model and found to have a 0.000001 parameter estimate which was found to be statistically significant. In this case the model is indicating that it is quite certain that the number of Tim Horton's has a very small impact on cost.