Appendix F. Description of Monte Carlo Model

Appendix F. Description of Monte Carlo Model

The IPA generated frequency distributions of annual electricity costs using a simulation model. The simulation model was implemented in Microsoft Excel using the @Risk add-in from Palisade Corp. (www.palisade.com). @Risk allows one to define certain parameters of a spreadsheet model as "random" and others as "outputs." @Risk will then recalculate the spreadsheet many times. Each recalculation ("iteration") uses a "sample" value for each random parameter. The samples are chosen to have specific distributional properties. For each output this process yields a large set of values – one for each iteration – and @Risk can report the entire set, key statistics such mean, variance and percentile points, or both.

This appendix describes first, the underlying spreadsheet model – the computation of annual cost – and second, the distributional assumptions used for random sampling. Some of the parameters used in the random sampling were assessed from historical data but others are "stylized facts" or reasonable assumptions. Therefore the results of the model are not intended as precise forecasts of costs but as indicators of which strategies are more or less risky.

F.1 Underlying Spreadsheet Model

This section describes the underlying spreadsheet model using algebraic notation. The model is for a single utility's portfolio (ComEd or Ameren). The following table summarizes all of the notation for variables and parameters.

Table of Notation

Today	is the date on which the model's forecasting is run
m	the index used for months in the model. There are twelve months representing a single delivery year from June to the following May. As a date, *m* is the first day of the month (used to calculate the days between the purchase of a forward contract, or the current forecast date, and the beginning of the delivery month)
p	a period within the month, either Peak or Off-peak
j	indexes the "legacy" forward contracts, that is, the previous IPA procurements for the subject utility
k	indexes the *K* separate procurements that will be used between the forecast date and the delivery to buy forward contracts. As a date, *k* is the date of the procurement
k+45	is the earliest a delivery month could begin for procurement *k* (45 days after the date of *k*) because procurements occur on April 15 (first contract is June) and September 15 (first contract is November)
sc	is a load forecast scenario (expected, high or low)
exp	is specifically the expected load forecast scenario
TotCost	is the total energy cost for the year

AvgLoad(sc, m, p)	is the forecast of scenario *sc* for the average load in MW in period *p* (Peak or Off-peak) in month *m*
LoadFcst(k,sc,m,p)	is the forecast of the average load in MW in period *p* (Peak or Off-peak) in month *m*, forecasted as of date *k*, based on partial information of what the load actually will be, assuming that the actual load will be *AvgLoad(sc,m,p)*
Hrs(m, p)	is the total number of hours in period *p* in month *m*
FwdMW(*j, m, p*)	is the megawatt volume of "legacy" forward contract *j* in period *p* (Peak or Off-peak) in month *m*
FwdMW(*k, m, p*)	is the megawatt volume of new forward contracts from procurement event *k* in period *p* (Peak or Off-peak) in month *m*
Tgt(k,m,p)	is the target (as a percentage of load) for the amount to be hedged up to and including procurement event *k*
Gran	is the forward contract granularity (size in MW) for procurement
ResidMWh(*m, p*)	is the residual *energy* volume in megawatthours – the amount of energy that must be purchased from the spot market -- in period *p* (Peak or Off-peak) in month *m*
FwdPri(*j, m, p*)	is the forward price for period *p* (Peak or Off-peak) in month *m* according to "legacy" forward contract *j*
FwdPri(*k, m, p*)	is the forward price for period *p* (Peak or Off-peak) in month *m* according to new forward contract *j* – the forward curve on date *k*
SpotPri(*m, p*)	is the average spot price for period *p* (Peak or Off-peak) in month *m*
ShpdPri(*m, p*)	is the shaped spot price for period *p* (Peak or Off-peak) in month *m.* Shaped means load-weighted according to *total* load
ResidPri(*m, p*)	is the "residual" spot price for period *p* (Peak or Off-peak) in month *m.* It is the average price paid for "residual load", which is the load not met by forward contracts
CVPri(m,p)	is a historical estimate of the "coefficient of variation" (ratio between population standard deviation and population mean) of PJM Northern Illinois day-ahead prices within period *p* (Peak or Off-peak) of calendar month *m* (in other words, if *m* is July 2015 then the estimate is over a set of historical Julys)
CVLd(m,p)	is a historical estimate of the "coefficient of variation" (ratio between population standard deviation and population mean) of hourly ComEd loads within period *p* (Peak or Off-peak) of calendar month *m* (in other words, if *m* is July 2015 then the estimate is

	over a set of historical Julys)
Corr(m,p)	is a historical estimate of the "correlation coefficient" (ratio between population covariance and the product of the two population standard deviations) of hourly PJM Northern Illinois day-ahead prices and ComEd loads within period *p* (Peak or Off-peak) of calendar month *m* (in other words, if *m* is July 2015 then the estimate over a set of historical Julys)

The following equations describe the model itself:

$$ResidMWh(m,p) = \left(AvgLoad(sc,m,p) - \sum_{j} FwdMW(j,m,p) - \sum_{k} FwdMW(k,m,p) \right) \cdot Hrs(m,p) \quad (1)$$

$$\begin{aligned} TotCost = & \sum_{m,j,p} FwdMW(j,m,p) \cdot Hrs(m,p) \cdot FwdPri(j,m,p) \\ & + \sum_{m,k,p} FwdMW(k,m,p) \cdot Hrs(m,p) \cdot FwdPri(k,m,p) \\ & + \sum_{m,p} ResidMWh(m,p) \cdot Hrs(m,p) \cdot ResidPri(m,p) \end{aligned} \quad (2)$$

The computation of ResidPri is complex and it is much easier to consider the forwards as if they were "contracts for differences":

$$\begin{aligned} TotCost = & \sum_{m,j,p} FwdMW(j,m,p) \cdot Hrs(m,p) \cdot \left(FwdPri(j,m,p) - SpotPri(m,p) \right) \\ & + \sum_{m,k,p} FwdMW(k,m,p) \cdot Hrs(m,p) \cdot \left(FwdPri(k,m,p) - SpotPri(m,p) \right) \\ & + \sum_{m,p} AvgLoad(sc,m,p) \cdot Hrs(m,p) \cdot ShpdPri(m,p) \end{aligned} \quad (3)$$

If there were a formula specified for *ShpdPri*, it would be exactly the formula needed to make (2) and (3) produce the same *TotCost*.

The model assumes that each procurement event has a target amount of forward energy to buy for each month, and that the IPA succeeds in buying the exact targets. The target is a percentage of the total load, based on a load forecast at the point of procurement. The forecast at the point of procurement is based on the expected load, with a certain amount of "learning" of what the load actually will be:

$$LoadFcst(k,sc,m,p) = \frac{m-k}{m-Today} \cdot AvgLoad(ex,m,p) + \frac{k-Today}{m-Today} \cdot AvgLoad(sc,m,p) \quad (4)$$

For a procurement whose date is later – closer to *m* – the load forecast is closer to *AvgLoad(sc,m,p)* than to *AvgLoad(ex,m,p)*. The model simulates procurements on April 15 or on September 15. When the procurement date *k* is September 15 of the delivery year, it will actually be later than some of the first months for which procurement is modeled. In those cases equation (4) may produce counter-intuitive results but there is no actual harm because the model will not simulate procurement of forward contracts for months beginning earlier than k+45.

The amount to be procured in event k is the target as a percentage of the forecast not yet covered, rounded to the nearest whole number of contracts (nearest multiple of *Gran*). However, in general sales are not allowed so the amount is not permitted to be negative.

(5)
$$FwdMW(k,m,p) = Max\left(0, Round\left(Tgt(k,m,p)\cdot \left(AvgLoad(sc,m,p) - \sum_j FwdMW(j,m,p) - \sum_{k'<k} FwdMW(k',m,p)\right), Gran\right)\right)$$

In some cases sales are allowed, though (namely for evaluating the risk impact of forward rebalancing); for those events,

(6)
$$FwdMW(k,m,p) = Round\left(Tgt(k,m,p)\cdot \left(AvgLoad(sc,m,p) - \sum_j FwdMW(j,m,p) - \sum_{k'<k} FwdMW(k',m,p)\right), Gran\right)$$

If $m < k$+45, the amount to be procured is zero, that is, equation (5) or (6) is replaced by *FwdMW(k, m, p)* = 0.

Finally, the shaped price is estimated using historic PJM Northern Illinois day-ahead prices and ComEd loads, and applied to both ComEd and Ameren as corresponding MISO/Ameren data was not available:

(7) $$ShpdPri(m,p) = SpotPri(m,p)\cdot\left(1 + Corr(m,p)\cdot CVPri(m,p)\cdot CVLd(m,p)\right)$$

F.2 Probability Model

The probability model uses the following additional notation for random variables and parameters.

Table of Notation (2)

Y	Number of years from *Today* to the first month in the delivery year being forecasted
TrHist(t,y,m)	An array of 7 observations of the ratio between historical forward prices for the calendar month corresponding to month *m*, and the same prices *y* years earlier, indexed by *t* =1 to 7 and *y* = 1 to 3.
TrAvg(y,m)	The average of the 7 values *TrHist(1,y,m), …, TrHist(7,y,m).*
Tr#	A random integer from 1 to 7 chosen from a uniform discrete distribution (equal probabilities)
sc#	A random number, chosen to be either -1, 0, or 1, corresponding to the low, expected and high load scenarios respectively

σP	Annualized volatility of forward prices during the prompt month (the month before delivery)
σ1	Annualized volatility of forward prices during the month before the prompt month
σMult	Monthly degradation in annualized volatility of forward prices
σSpot	Standard deviation of the logarithm of the ratio of the monthly average spot price to the price for the forward contract for that month on the last day of the prompt month
LnMult(m, d1, d2)	Logarithm of the volatility-based perturbation in forward prices for month *m* from date *d1* to date *d2*, both of which are prior to the prompt month.
LnSpot(m, d1)	Logarithm of the volatility-based perturbation in spot prices for month *m* relative to forward prices observed prior to the prompt month, on date *d1*.

The fundamental random variables are load and price. The load variable is completely described by the scenario, *sc*. Price is derived from the current forward curve with the addition of a "trend" and volatility. The trend is a ratio and is based on observations, over the last seven years, of the ratio between the average of the PJM-W yearly day-ahead price strip and the same average observed one, two or three years earlier (depending on which year the procurement is for).

The trend is assumed to have some correlation with the load scenario – higher values of the trend should indicated that the high scenario is more likely to occur. The model uses a value of 20% as the correlation between *Tr#* and *sc#*. The seven values of *Tr#* are assumed occur with equal probabilities. For *sc#*, the values -1 and 1 are associated with extreme low and high cases and have less probability weight than the expected case (sc#=0); although -1 and +1 have the same weight. The computation of the weights on -1 and +1 is based on an assumption about the percentile points represented by the high and low scenarios as described at the end of section 3.5.2. The specific probabilities used for each utility are given in Table F-1.

Table F-1 Scenario probabilities

		Probability weights	
Scenario	*sc#*	Ameren	ComEd
Low	-1	0.1302	0.1848
Expected	0	0.7396	0.6304
High	1	0.1302	0.1848

The prices are assumed to "evolve" from the forward curve. In other words, the forward prices changes by some fraction of the trend, just as the load forecast evolves in equation (4). In addition the forward price experiences random fluctuations, associated with volatility. The volatility of the price is an annualized number. If the annualized volatility of price is σ, then the change in price over D days is a lognormal random variable with mean 1, and the logarithm of the change has standard deviation $\sigma\sqrt{D/365}$.

The volatility of forward prices grows as the contract gets closer to delivery. Values of *σP*, *σ1*, and *σMult* were assessed from historical data. *σP* is the annualized volatility of forward prices during the month prior to delivery. *σ1* is the annualized volatility of forward prices during the second month prior to delivery (the month before the prompt month); *σ1* < *σP*. The annualized volatility of forward prices is then *σMult·σ1* during the second month prior to delivery, $\sigma Mult^2 \cdot \sigma 1$ during the second month prior to delivery, etc. (*σMult* < 1).

The volatility model says that the change in forward prices from date *d1* to date *d2* (excluding any price trend), as long as both are before the start of the prompt month, is a lognormally distributed random variable denoted *LnMult(m, d1, d2)*. Its variance depends on *σ1, σMult, d1,* and *d2*, and its expected value is -½ the variance (which is necessary to make the expectation of $e^{LnMult(m,\, d1,\, d2)}$ equal to 1). The ratio between the forward price for month *m* observed on date *d1* and the average spot price for that month is also a lognormally distributed random variable denoted *LnSpot(m, d1),* whose variance depends on *σP* and *σSpot* as well as *σ1, σMult, m,*and *d2* and its expected value is -½ the variance.

Recall that the model assumes K separate procurements, their dates may be sequenced as $k_1, k_2, \ldots k_K$. The underlying random variables are $LnMult(m, Today, k_1)$, $LnMult(m, k_1, k_2)$, ..., $LnMult(m, k_{K-1}, k_K)$, and $LnSpot(m, k_K)$, and the forward prices (including random trend and random volatility-based perturbations) are

$$FwdPri(k_1, m, p) = FwdCrv(m, p) \cdot \left(\left(TrHist(Tr\#, Y, m) - TrAvg(Y, m)\right) \cdot \frac{k_1 - Today}{m - Today} \right) \cdot e^{LnMult(m, Today, k_1)}$$

$$FwdPri(k_2, m, p) = FwdPri(k_1, m, p) \cdot \left(\frac{k_2 - Today}{k_1 - Today} \right) \cdot e^{LnMult(m, k_1, k_2)}$$

$$\vdots$$

$$SpotPri(m, p) = FwdPri(k_K, m, p) \cdot \left(\frac{m - Today}{k_K - Today} \right) \cdot e^{LnSpot(m, k_K)}$$