



A GENERATION AHEAD,
today

**Comments on PJM & ITRON
Load Forecasting Recommendations
*Draft***

Date: 12/6/2010



Summary



- Overall, we prefer the Average GMP Model (GMP) for its simplicity. We suggest that more research be done before PJM adopts any new methodology.
- The Average Index1 Model (Index1) improvements in statistical fit are marginal at best. It is not systematically better than GMP model across all the PJM regions.
- Index1 has a systematic downward bias because the higher weights are assigned slower growth variables- Population and Households.
- The weights in Index1 should be based on the marginal contribution to explain the load growth from each economic series not on a survey.
- Index1 seems to violate the *Principle of Parsimony - Simple explanations should be preferred to complicated ones.*
- Load analysis and simulation should reflect not only the weather risks but also the volatility in the economic forecasts.
- We believe that there are some serious theoretical flaws and shortcomings in the composition of Index1.
 - The static weights, in our opinion, are highly subjective and even arbitrary from a statistical point of view.
 - The weights play a vital role in determining the economic index and should be estimated from historical dataset not from surveys.

Index1 is not a clear winner



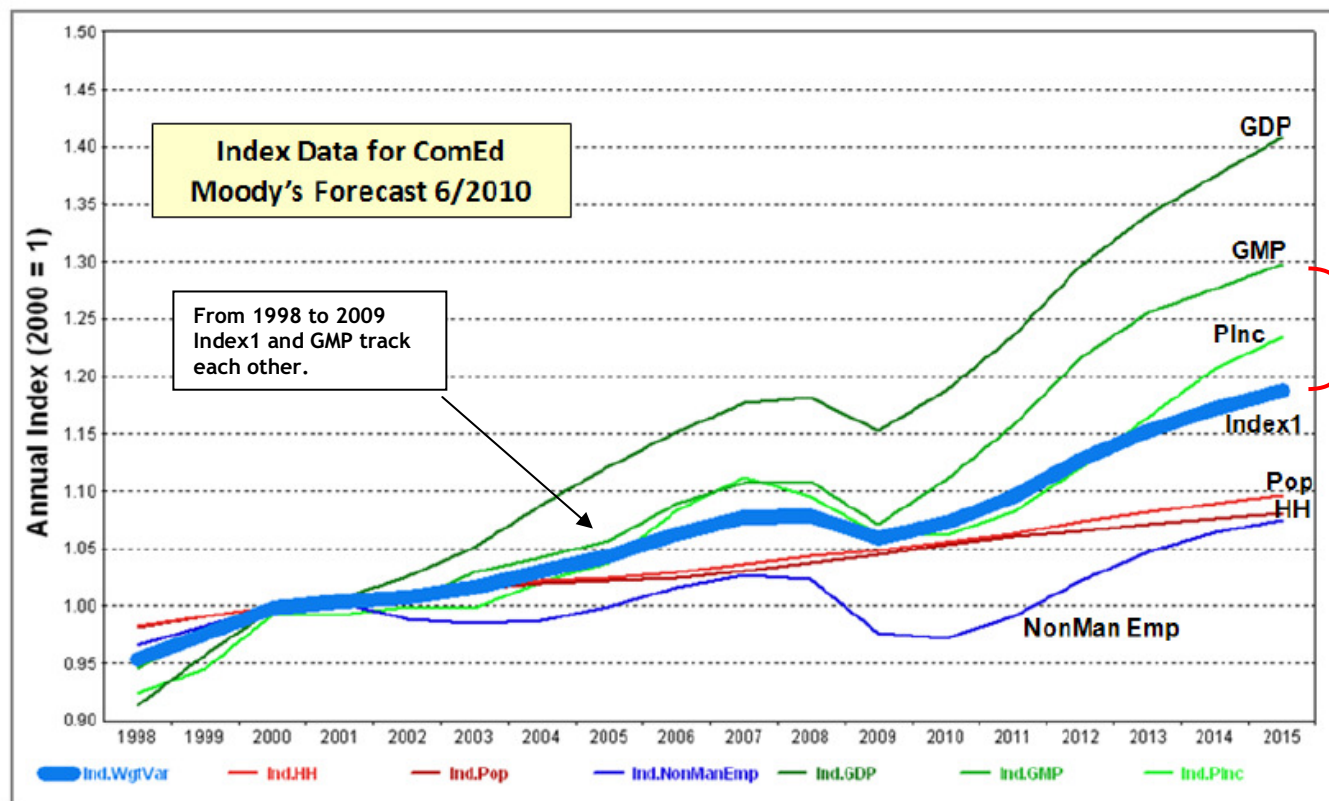
- When Index1 fits better than the GMP model, the improvement is marginal.
 - MAPE improvement for most zones is less than 0.1%.
 - Adjusted R^2 improvement for most zones is less than 0.5%.
- We question the high R^2
 - High R^2 with low Durbin-Watson statistic (a measure of autocorrelation) is the symptom of a spurious regression¹.
 - We are interested in seeing the residuals from both the forecast methodologies, as well as, the autocorrelations of those residuals.
 - Autocorrelation in the residuals can mean good in-sample fit but poor forecasting performance.
- Looking at MAPE, T-stats and R-square are not sufficient to compare the forecasting accuracy of the two models.
 - In-sample statistics tell about the past performance. Out-of-sample statistics posted on 12/2/2010 rely on the current vintage of economic series which have the most-recent estimated realized values for the past (before 2009?) and forecast for 2010 and beyond. Thus, they only test the model specification but not the accuracy of each forecasted economic series.

1. Grainger & Newbold, "Spurious Regression In Econometrics", Journal of Econometrics, 1974

Index1 has a systematic downward bias because the higher weights are assigned slower growth variables



Figure 29: Weighted Economic Variable (Commonwealth Edison)



Source: ITRON report p. 29: 20101006-item-09-load-forecasting-recommendations-phase-1-findings

$$\text{Index } 1_{y,m} = \left(\frac{\text{HH}_{y,m}}{\text{HH}_{\text{base}}}\right)^{0.28} \times \left(\frac{\text{Pop}_{y,m}}{\text{Pop}_{\text{base}}}\right)^{0.11} \times \left(\frac{\text{NMEmp}_{y,m}}{\text{NMEmp}_{\text{base}}}\right)^{0.15} \times \left(\frac{\text{GDP}_{y,m}}{\text{GDP}_{\text{base}}}\right)^{0.17} \times \left(\frac{\text{GMP}_{y,m}}{\text{GMP}_{\text{base}}}\right)^{0.15} \times \left(\frac{\text{Pinc}_{y,m}}{\text{Pinc}_{\text{base}}}\right)^{0.14}$$

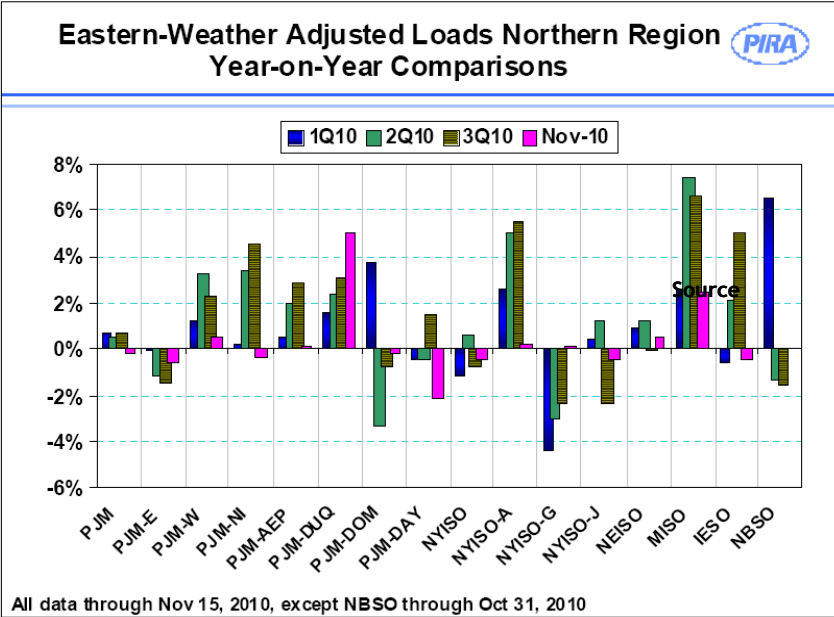
HH-households, **Pop** - Population, **NMEmp** - non-manufacturing employment, **GDP** - gross domestic product, **GMP** - gross metropolitan product, **Pinc**- personal income

Draft - The information contained in this presentation is not meant to provide forward-looking information with respect to Calpine's financial performance.

Index1 puts too much weight on slow growth variables such as Population and Households



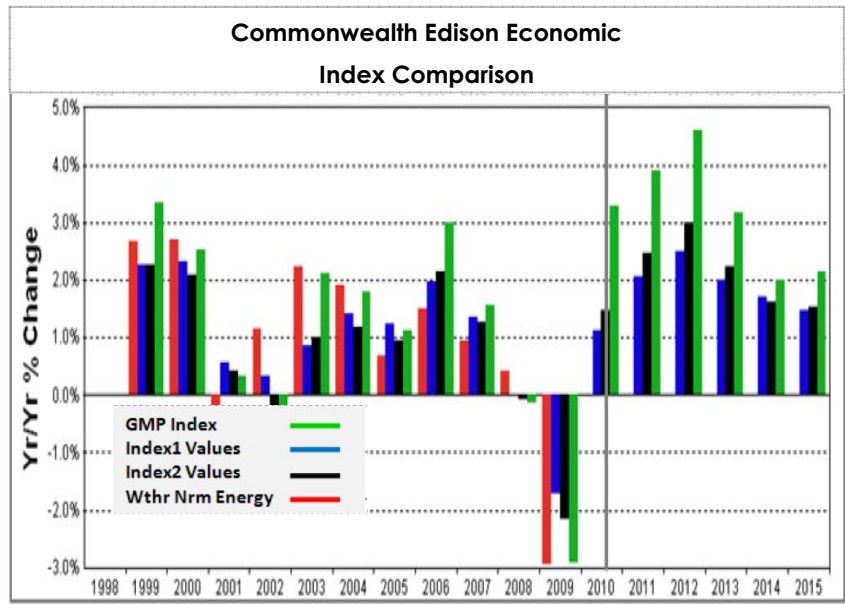
Load changes are much more sensitive to GMP and GDP than to other variables.



Source: PIRA

- Based on PIRA estimates, the changes have been above 4% for Commonwealth Edison during Q3-2010.
- Weather-adjusted load changes are much larger than those predicted by Index1. (see chart top right)
- Load changes are much more sensitive to Gross Metropolitan Product and GDP than to other variables.

Actual load follows GMP more closely than Index1 during 2009 & 2010



Source: ITRON Review of PJM Load Forecasting Models Phase I, p. 44

- GMP Model captures the large downward and upward year-on-year load changes in 2009 and 2010, respectively.
- Population and households have low and steady growth rate, so Index1 reduces load growth relative to GMP.
- The weights for population, household and non-manufacturing employment should be lower versus GMP & GDP.

Draft Version

The Index1 weights should be estimated from historical dataset not from surveys



- The Index1 weights are set based on survey
 - They are the same for each load zone despite different economic drivers in the load zone.
 - They are constant over time with no reflection on the changing economy.
 - The weights determine the contribution to the exponential growth rate of Index1 from each component's exponential growth rate.
 - We are surprised that Gross Metropolitan Product has been, thus far, the sole economic driver; but its weight has now been reduced to 15%.
- Since Index1 is used for forecasting the peak load, the weight should be based on the marginal contribution to explain the load growth from each economic series.
 - Principal Component Analysis (PCA) can deal with the multicollinearity of the six economic series statistically estimated and compare the weights for Index1
 - PCA could also be applied to the economic series to formulate one or two unique economics series
- We question the weights being the same across all the regions since each region's economy and makeup are different.
- Index 1, uses several economic variables that have dissimilar sampling frequencies; which introduces additional forecasting error.

Index1 seems to violate the *Principle of Parsimony*¹ - Simple explanations should preferred to complicated ones



- ***Models should have as few parameters as possible.***
 - Our view is that Index1 has redundant variables and at minimum could be made simpler.
 - Population \approx Households x Avg. Persons/Household
 - GDP & Gross Metropolitan Product both measure the same thing but it seems like Gross Metropolitan Product is more appropriate because of its local nature, i.e., strong GDP growth probably does not mean anything to PJM Regional load unless it implies Gross Metropolitan Product growth.
 - We suspect Non-Manufacturing Employment, GDP, Gross Metropolitan Product, and Personal Income are cross correlated and serially correlated and the index could likely be simplified.
 - ITRON started with a large number of variables but they did not determine the *minimal adequate* model.
- ***Experiments relying on few assumptions should be preferred to those relying on many.***
 - Index1 requires forecasts for *six* series to be accurate. The GMP relies on *five less* variables.
- ***Linear models should be preferred to nonlinear models.***
 - Index1 is a *geometrically* weighted average of six series. GMP is simply multiplied by single factor.
 - Index1 is more difficult to interpret - how do you answer the question, “How much does the peak load change with a 1% GDP change?”

¹ The Principle of Parsimony can be found in *The R Book* by Michael Crawley

Load simulation should reflect GMP forecast volatility



- As we've witnessed the large swing in GMP forecasts lately and the constantly shifting and diverging views of economists about the state of economy, load simulation should include scenarios of different economic growth assumptions in addition to historical weathers.
- The new simulation results with GMP forecast volatility will improve the accuracy of 90/10 peak load forecast, IRM analysis and CETO study.
- GMP forecast volatility may also need to be a part of the 50/50 peak load forecast. The base GMP forecast can be at the upper range of historical forecasts or can be the current forecast plus one standard deviation of historical forecasts during last two years, for example.

