STATE OF ILLINOIS

ILLINOIS COMMERCE COMMISSION

NORTHERN ILLINOIS GAS COMPANY D/B/A NICOR GAS COMPANY	:	No. 04-_____
	:	
Proposed general increase in rates, and revisions to other terms and conditions of service.	:	

Direct Testimony of

CESAR A. HERRERA, PH.D.

Senior Consultant

National Economic Research Associates, Inc.

on behalf of

Northern Illinois Gas Company

Nicor Gas Exhibit 15.0

Q. Please state your name and business address.

A. My name is Cesar A. Herrera. My business address is 200 Clarendon Street, 35th floor, Boston, Massachusetts, 02116.

Q. How are you employed?

A. I am a Senior Consultant at National Economic Research Associates Incorporated (NERA).

Q. Please describe your academic background.

A. I have a B.Soc.Sc. in Economics from the Catholic University of Peru, an M.A. in Development Economics from the University of Sussex, England, and an M.Sc. and Ph.D. in Economics from the University of Wisconsin-Madison.

Q. Please summarize your work experience.

A. I specialize in tariff regulation and the economics of utility businesses, including competitive restructuring issues. I have participated in a number of studies of the electric, natural gas and railroad industries, both domestically and abroad. My consulting and research work involves applied microeconomics and statistics with a focus on regulated activities. I have assisted clients with a variety of empirical issues including productivity and cost performance measurement, valuation of energy contracts, assessment of economic damages, statistical support in tariff reviews, and the modeling of wholesale energy markets. My current C.V. is attached to this testimony as **Exhibit 15.1**.

Q. Dr. Herrera, what is the purpose of your testimony?

A. I will address three quantitative issues that relate to billing determinants. First, I will explain how I obtained a forecast of heating degree days ("HDD") for Northern Illinois

Gas Company (to which I refer in my testimony as Nicor Gas or simply the Company) for 2005; second, I will explain how I derived a normalized daily weather (HDD) profile for the test year based on this forecast; third, I will discuss the minimum sample size needed to estimate accurately the percentage of the company's uncollectible charge-offs that are attributable to the commodity component.

Forecast of Heating Degree Days

Q. On what did you base your forecast of HDD?

A. I based the forecast on actual historical observations.

Q. Where did you obtain your data?

A. I received from Nicor Gas the yearly HDD data recorded at the Midway Airport weather station, from 1929 to 2003. The actual data used is shown in **Exhibit 15.2**.

Q. Please explain how you used this data to obtain a HDD forecast for 2005.

A. I used a common forecasting technique that estimates the average yearly HDD for a given time frame, and then uses these results to predict weather in the forecast year. For this analysis, I tested two alternative means of forecasting HDD in 2005: 1) a 30-year average of HDD data ending in 2003, and 2) a 10-year average of HDD data also ending in 2003. I then conducted a statistical comparison of the predictive capability of these two time horizons to determine which was more appropriate. Specifically, I first calculated and compared the root mean squared error for each of the two averaging periods. Second, to better understand these, I used a linear regression technique to examine the trending behavior of the HDD data.

Q. Please describe how you analyzed the HDD data.

A. The data series from Midway Airport weather station begins in 1929, so it was possible to calculate both 10-year and 30-year averages for the years 1958 through 2003. I compared the 10-year and 30-year average HDD figure for each year with the actual temperature observed 2 years later. For example, I compared the 10-year and 30-year averages for 1958 with the actual temperature for 1960, recording the difference (error) between the actual and forecasted values for each. I repeated this process to 2001—the most recent year for which actual data existed two years later. The analysis I conducted parallels the situation with which Nicor Gas is confronted: using HDD data of the most recent year available (2003) in order to predict weather two years ahead (2005).

Q. How did you compare the predictive capabilities of the two averages?

A. I conducted a statistical analysis to compare the predictive capabilities of the 10-year and 30-year averages. I calculated a standard statistic called the "root mean squared error" (RMSE). The RMSE statistic, which is widely used to measure the accuracy of forecasts, is a number representing the degree to which the forecasted values fail to correspond to the actual data. The smaller the RMSE, the smaller the overall difference between the actual and forecasted HDD. The formula for the RMSE is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(HDD_i - HDD_i^F\right)^2}$$

where the letter i denotes year of the observation, n denotes the total number of years (*i.e.*, 44), HDD_i refers to actual values, and HDD_i^F is the forecasted HDD. $\left(HDD_i - HDD_i^F\right)$, therefore, measures the difference between actual and forecasted value.

Q. Please describe your results.

A. My results are shown in **Table 1**. Based on Midway weather station's historical data, a 10-year HDD average outperforms a 30-year average in predicting weather two years into the future. In particular, as a forecasting instrument, 10-year averages tend to produce more precise forecasts than 30-year averages. This statistic shows that the forecast errors of 30-year averages are typically higher than those of 10-year averages, by a magnitude of about 8 HDD. Based on the RMSE test, therefore, a 10-year average represents a better option for purposes of forecasting HDD in 2005. The detailed results of this analysis are in **Exhibit 15.3**.

Table 1. RMSE of Predicting Weather on the Basis of Historic HDD Averages

HDD Average	RMSE
30-year	453.81
10-year	445.95

Q. You said you examined the data using a linear regression technique. What is that?

A. Linear regression is a technique used to explain the relationship between two variables by finding a straight line that best fits the data. Here the two variables are the year and the observed HDD, and the linear regression technique can be used to estimate a time trend in the Midway HDD data. I then used this model to obtain a HDD forecast for 2005, using the trend line. To perform the regressions, I used STATA, a commonly used statistical software package.

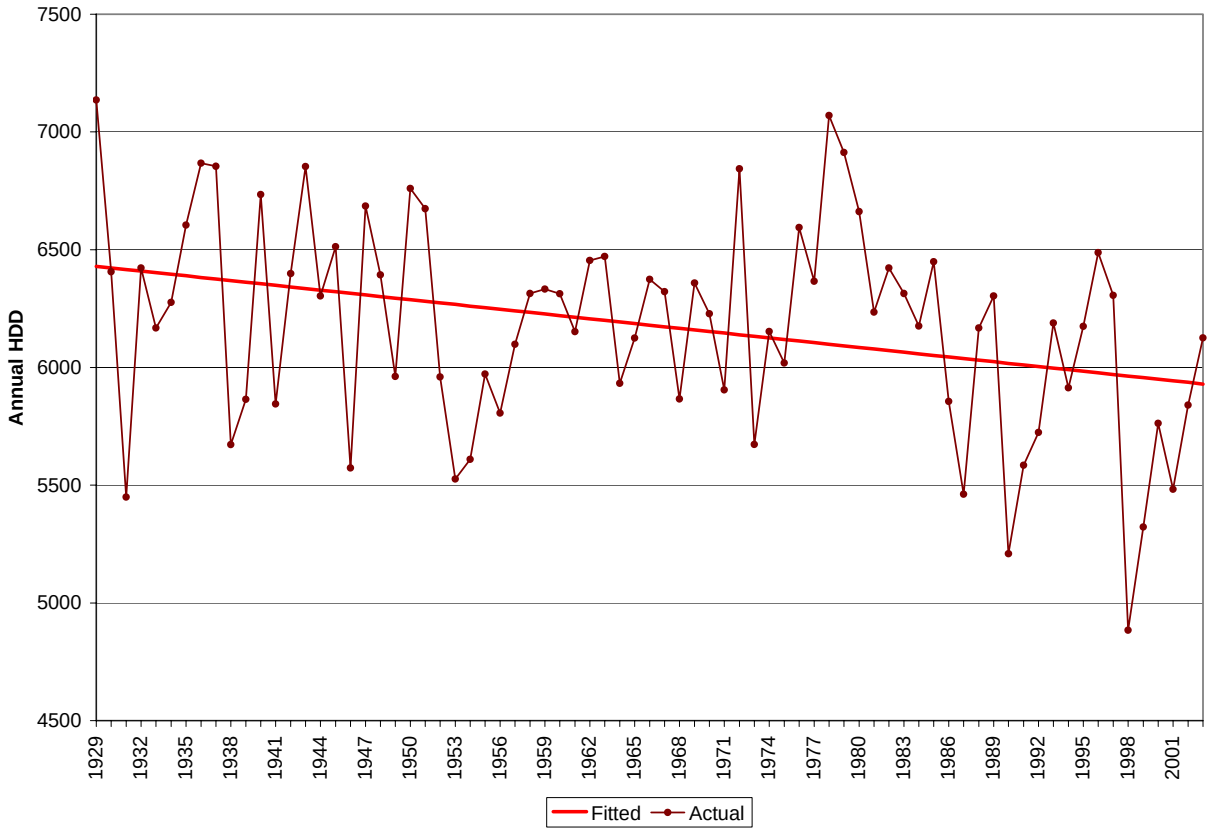
Q. What time period of data did you use to arrive at a 2005 forecast?

A. I used HDD data from Midway for the years 1929-2003. As explained below, I also used data for two sub-periods: 1929-1975 and 1976-2003.

Q. What were the results of your regression analysis?

A. **Figure 1** shows the annual HDD observations and the trend line that best fits these points for the years 1929-2003. The fitted line shows a downward trend that is statistically significant as measured by a t-statistic (see result details in **Exhibit 15.4**).

Figure 1. Actual and Fitted Annual HDD for Midway Weather Station, 1929-2003



Q. What is the importance of this trend line to your analysis?

A. If a data series is "trend-less" (*i.e.* the line slope is equal to zero) then its mean value will remain stable in time. In this case, an average calculated over a wide set of observations would be a good predictor of future values given that data is essentially (mean) stationary. However, the Midway HDD data shows a significant downward trend. Any

forecast that relies on an average taken over a long period of time (*e.g.*, 30 years) runs the risk of not reflecting the expected evolution of the HDD series as accurately as an average based on a shorter time period (*e.g.*, 10 years), particularly if the average refers to a more recent period.

Q. The trend line is a straight line for the entire period. Does this mean that the trend would be the same regardless of the range of years?

A. Not necessarily. Indeed, based on visual inspection of the Midway HDD data, there appears to be some type of inflection, or shift, in the slope of the time trend, perhaps at some point in the 1970s.

Q. Are there statistical tests to establish whether such a shift did actually occur?

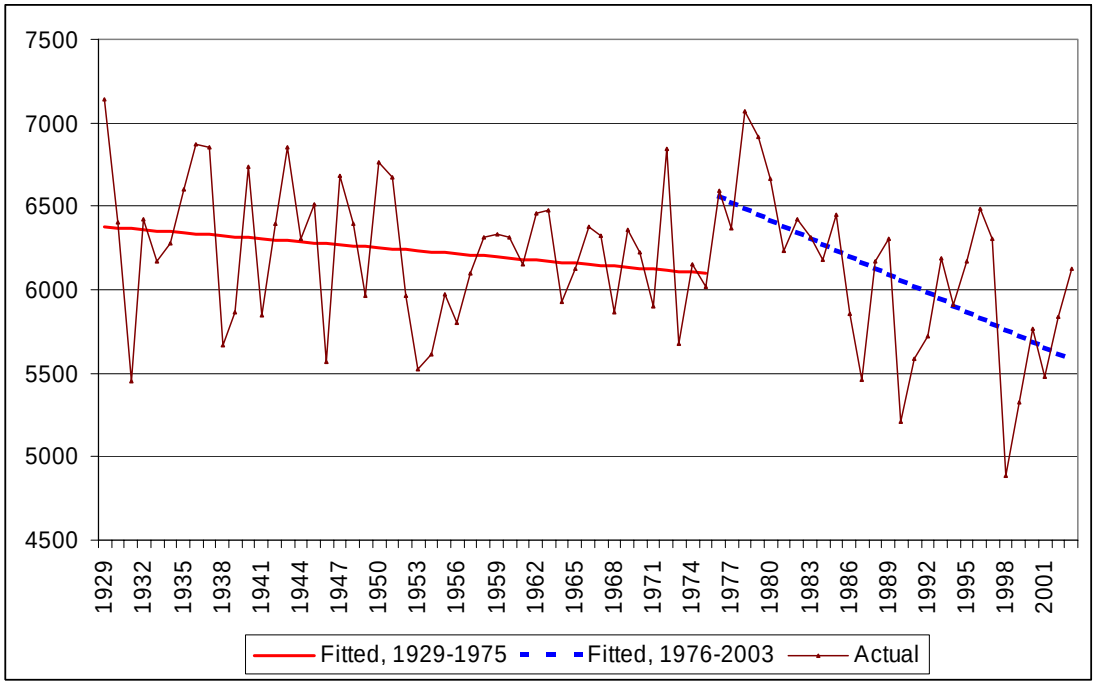
A. Yes. There are statistical tests that can help to establish the existence of such a trend break and the date when it occurred. I performed several of these tests, which indicate that there was a structural break in the HDD series in 1975 (see **Exhibit 15.5** for a description of these tests). Following this result, I applied the regression model to estimate two trend lines, one up to the break date and the other after the break date.

Q. What was the result of your analysis?

A. **Figure 2** shows the annual HDD observations from 1929-2003, and two distinct trend lines—one that best fits the 1929-1975 data points, and another that best fits the 1976-2003 data points. The statistical results show that the slope in the earlier trend line is also negative but not statistically significant, whereas the coefficient of the more recent trend line is both negative and statistically significant (see details in **Exhibit 15.6**). In other words, until the mid-1970s, there appears to be no discernable trend in the behavior of

Midway HDD, but after that point, there is a significant downward trend (*i.e.*, lower HDD, or warmer winters).

Figure 2. Fitted HDD Trend Lines for Two Sub Periods, 1929-1975 and 1976-2003

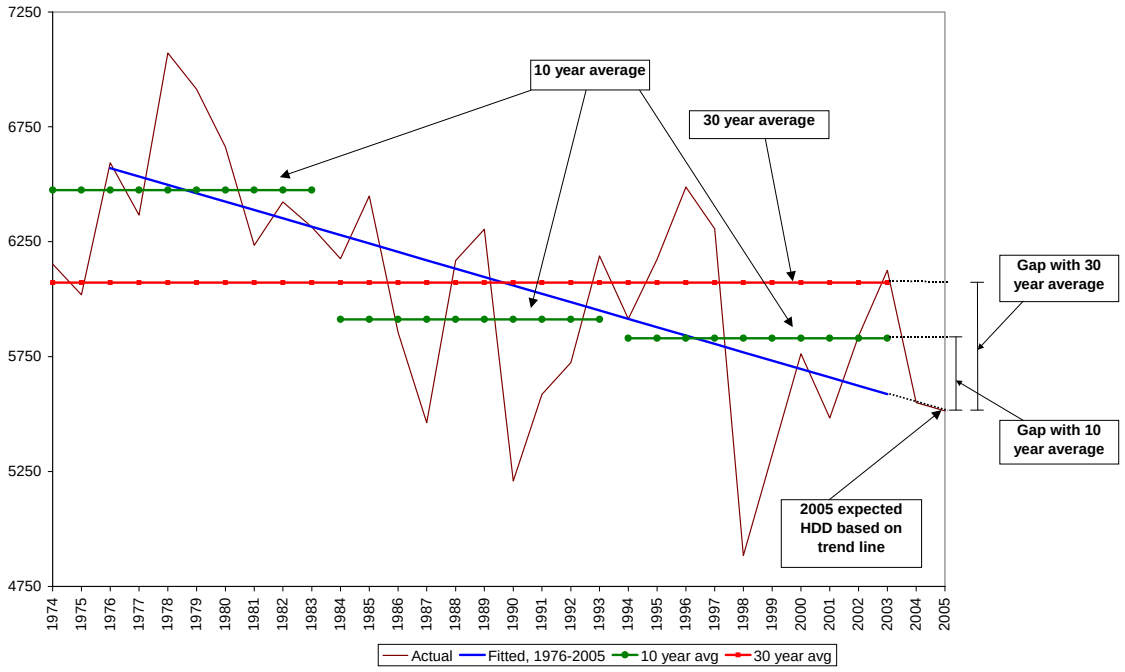


Q. What is the importance of the trend line after the break, the 28-year trend, to your analysis?

A. **Figure 3** shows the Midway HDD data for the sub-period 1974-2003 together with the fitted trend line following the 1975 break. It also shows the HDD levels that correspond to the 30-year average and 10-year averages that cover the three-decade period. **Figure 3** illustrates why a 10-year average does a relatively better job of predicting the expected behavior of the Midway HDD than a 30-year average—it is able to track the downward trend more closely. For instance, by extrapolating the fitted trend line, we can obtain a forecast of the 2005 HDD (see table below for an exact estimate). When compared with the 30-year and 10-year averages, we notice that the gap between the trend forecast and

the HDD average is noticeably wider in the 30-year case than with the 10-year. Therefore, with the significantly declining HDD trend a 30-year average will be more prone to overestimating (*i.e.*, producing less accurate) HDD forecasts than a 10-year average.

Figure 3. Midway HDD Actual, Fitted and Average Values, 1974-2003



Q. What do you conclude from your forecasting results?

A. Once I estimated the three regression models, my next step was to obtain HDD forecasts for 2005. For the 75-year and 28-year trend lines, the prediction is the number predicted by continuing the trend line two additional years. The predicted values are shown in **Table 2**. For comparison purposes, I also include the magnitudes corresponding to the 2003 30-year and 10-year HDD averages at Midway.

Table 2. 2005 Forecasts of Annual HDDs at Midway Weather Station

Source of Forecast	HDD
1929-2003 Trend Line (75 Years)	5916
1976-2003 Trend Line (28 Years)	5514
30-Year Average (as of 2003)	6072
10-Year Average (as of 2003)	5830

As noted above, given the clear downward trend in HDD, a 30-year weather average will tend to overstate expected HDD for 2005. A 10-average has less of a tendency to overstate HDD since it focuses on a more recent period, thus partly accounting for the downward trend in HDD. Finally, the regression-based predictions explicitly account for weather trending behavior, which indicates a significant decrease in HDD over time, particularly since the mid-1970s. Based on this observation and the RMSE test results, I conclude that a 10-year HDD average, 5830 HDD, provides the most appropriate forecast of typical weather in 2005.

Normalized Monthly HDD

Q. For the purpose of estimating peak monthly demand, Nicor Gas needs normalized monthly HDD statistics. What calculations did you perform to calculate the normalized monthly HDD based on the daily 10-year HDD data that you were presented?

A. I developed a set of "normal" values by "smoothing" a series of daily 10-year HDD averages from the Midway Airport weather station.

Q. Please describe the "smoothing" process you used to normalize HDD data for Nicor Gas.

A. The "smoothing" approach that I used involved the following steps:

- Summation of daily degree-day values into monthly HDD estimates. I used daily degree-day observations for the period January 1, 1994 through December 31, 2003 (10 years) as input information. For each calendar month (January-

December), I produced ten monthly degree day totals, one for each year in my sample;

- Averaging (over these 10 years) the total HDD for each month to produce a preliminary monthly normal degree-day;
- Interpolating the daily normal HDD values, which I performed on a monthly basis. I first obtained normal degree-days for each month by smoothing HDD data for that month, plus the two weeks preceding and following that month;
- Adjusting each month such that its monthly smoothed HDD total equaled the historical HDD 10-year average for that month. The degree of smoothing was determined by using a bandwidth of 0.4, which keeps the smoothed results relatively close to the actual data;
- Rounding the smoothed daily numbers to the nearest integer. I scaled the data so that monthly totals matched the 10-year averages, also to the nearest integer.

Q. What were the results of your analysis?

A. **Exhibit 15.7** shows the actual daily HDD 10-year averages at Midway and the monthly smoothed values that I obtained (both before and after rounding).

<u>Sample Size for Analyzing Uncollectible Accounts</u>

Q. Dr. Herrera, what was the purpose for sampling of uncollectible accounts?

A. The purpose of the sampling analysis is to determine the fraction of Nicor Gas' uncollectible account balances that correspond to gas supply costs. Nicor Gas has a total of over 111,000 customer accounts that were charged-off during a two-year period. However, calculating the commodity component of each bill is an extremely time consuming and costly manual process. In such situations, it is common practice to conduct a sampling analysis to estimate a given population parameter. With an appropriate sample, Nicor Gas can obtain a satisfactory estimation of the percentage of charge-off corresponding to commodity and non-commodity components.

Q. What were you asked to do?

A. The Company performed a stratified random sample with a sample size of 68 accounts from each of three classes: residential space heating, residential non-space heating, and non-residential. The Company asked me to review its sampling methodology using accepted, standard statistical formulas.

Q. What were your results?

A. I found that the Company's sample size is large enough to estimate the population gas supply cost to total charge-off ratio, using a proportionate stratified random sampling approach, to give a 95% confidence level and no more than a 5% margin of error.

Q. Does this complete your direct testimony?

A. Yes.