Forecasting Retail Prices in the Indian Cement Industry

by Soumen Karkun & Rahul Sadhu, Holtec Consulting, India

1. INTRODUCTION

Despite the demand for cement being forecast to grow at near double-digit rates, supply side capacity buildups are likely to create a situation of significant surplus, at the inception of the next decade. Concerns abound of whether such a surplus would depress retail prices, and if so, to what extent. With business planning as well as investment decisions suffering consequent uncertainty, this paper attempts to devise a methodology to forecast price.

The output of the exercise is, in the opinion of the authors, a fairly robust model that would absorb situational aberrations suffered by different variables, in different segments of the industry life cycle. While its application in India is currently under pilot testing, the authors feel that the basic model appears rugged enough to be applied in the cement industries of other countries as well.

In the context of this paper, FY 2008 refers to the period between 1 April 2007 - 31 March 2008. The authors present their conclusions on the retail prices of cement in India that are likely to prevail over the period FY 2009 - FY 2015.

2. METHODOLOGY

Unlike the multiple models existing in the petro-chemical industry, the authors were unable to find a comprehensive model for forecasting retail prices in the global cement industry. Consequently, the model presented in this paper had to be developed from scratch. The first step in the development process was the formulation of an imaginative list of “lead” variables that could possibly impact retail price. This required not only an understanding of econometrics, but also a thorough comprehension of how the industry actually worked. Other key steps in the development process included the determination of statistical distributions for these variables, the derivation of their causal inter-relationships and the application of a boutique of quantitative techniques (conventional statistics, curve fitting, exponential smoothing, correlation exercises, multiple regression, Monte Carlo simulation, etc.). Reality checks were diligently applied on all quantitative outcomes by cross-checking whether the model accurately estimated the past. This was achieved by applying industry know-how in formulating basic assumptions and rejecting spurious solutions.

The methodology is better illustrated by the following steps, which were performed, though not necessarily in the given order. While some steps ran in parallel, others were iteratively influenced:

1) Listing out all possible variables that could possibly impact cement prices.

2) Searching and selecting from secondary sources for past data on these variables.

3) Developing a shortlist of variables based on industry knowledge, lead behavior with respect to price, independence, and sustained data availability & reliability.

4) Modeling quantitative relationships between price and the shortlisted variables.

5) Predicting the future behaviour of these variables by examining, what, these in turn, depended upon.

6) Forecasting price in a “static” (deterministic) mode, based on the quantitative relationships modeled in Step 4) and the predicted values for the variables, as determined in Step 5).

7) Selecting the best model based on statistical tests applied to alternative models.

8) Conducting reality checks on the model to examine its accuracy of fit to actual price behaviour in the past as also to eliminate the chances of a spurious fit.

9) Postulating price ceilings (based on the elasticity of demand and "affordability") and floors (based on the future cost of production inputs, capital investments and financing costs).

10) Freezing the "static" model and developing forecasts price based on its application.

11) Developing statistical distributions for the shortlisted variables. Where no specific inputs were available, these were based on the variabilities observed in the past corrected by an anticipation of influencing events in the future.

12) Deriving the "dynamic" (stochastic) model for price by simulating the independent variables and thus producing a simulated output for the price. Proprietary simulation software (XLSim® from AnalyCorp, Inc.) was used for this purpose.

13) Analysing the simulated prices for each year in order to obtain the mean and the upper and lower 90% confidence bands.

3. VARIABLES

Initial brainstorming resulted in a fairly exhaustive list of 42 *basic* variables, classified into 3 major groups as shown below :

- **Economic/ Econometric Variables** : GDP & GDP Per Capita, Population Demographics, Inflation (WPI - All commodities), Construction Activity Index, Exchange Rate, Planning Cycles, Triggering Events (e.g. Election Year, etc.), Import Duties, Domestic Duties & Taxes, Unfulfilled Housing, Construction/ Cement Spend (Volumes x Price), etc.

- **Cement Industry Variables** : Demand & Demand Concentration, Unfulfilled/ Latent Demand, Per Capita Consumption, Price Elasticity of Demand, Price Inertia (effect that previous prices have on future prices) Capacity, Scale Economics in Capacity Creation, Supply & Supply Concentration, Imports & Import Prices, Exports & Export Prices, Capacity Utilisation, Degree of Fragmentation/ Industry Consolidation, End User Segmentation (e.g. % consumed by Individual Housing), % of Blended Cement in the Market, Availability & Costs of Inputs (e.g. Limestone, Blending Materials, Fuel, Power, etc), Technology Shifts, Expected Costs of Production, etc.

- **Financial & Cost Variables** : Expected Investment Returns, Share Price Indices, Cost of Debt, Prices of Complementary Products (e.g. Steel, etc), Prices in Real Estate (e.g. Land, Apartments, Commercial Buildings, etc), Freights, etc

Decision filters applied to the longlist to arrive at a shortlist included :

- Availability of **reliable data** about the variable in the past.

- **Degree of independence** that the variable enjoyed in relation to the others.

- **Extent of correlation** between the variable and the retail price of cement

- **Lead-lag relationship** between the variable and the retail price of cement.

- **Forecasting accuracy** for the variable.

- **Combinability** of the variable with others.

The sequential application of filters is depicted in **Figure 1**.

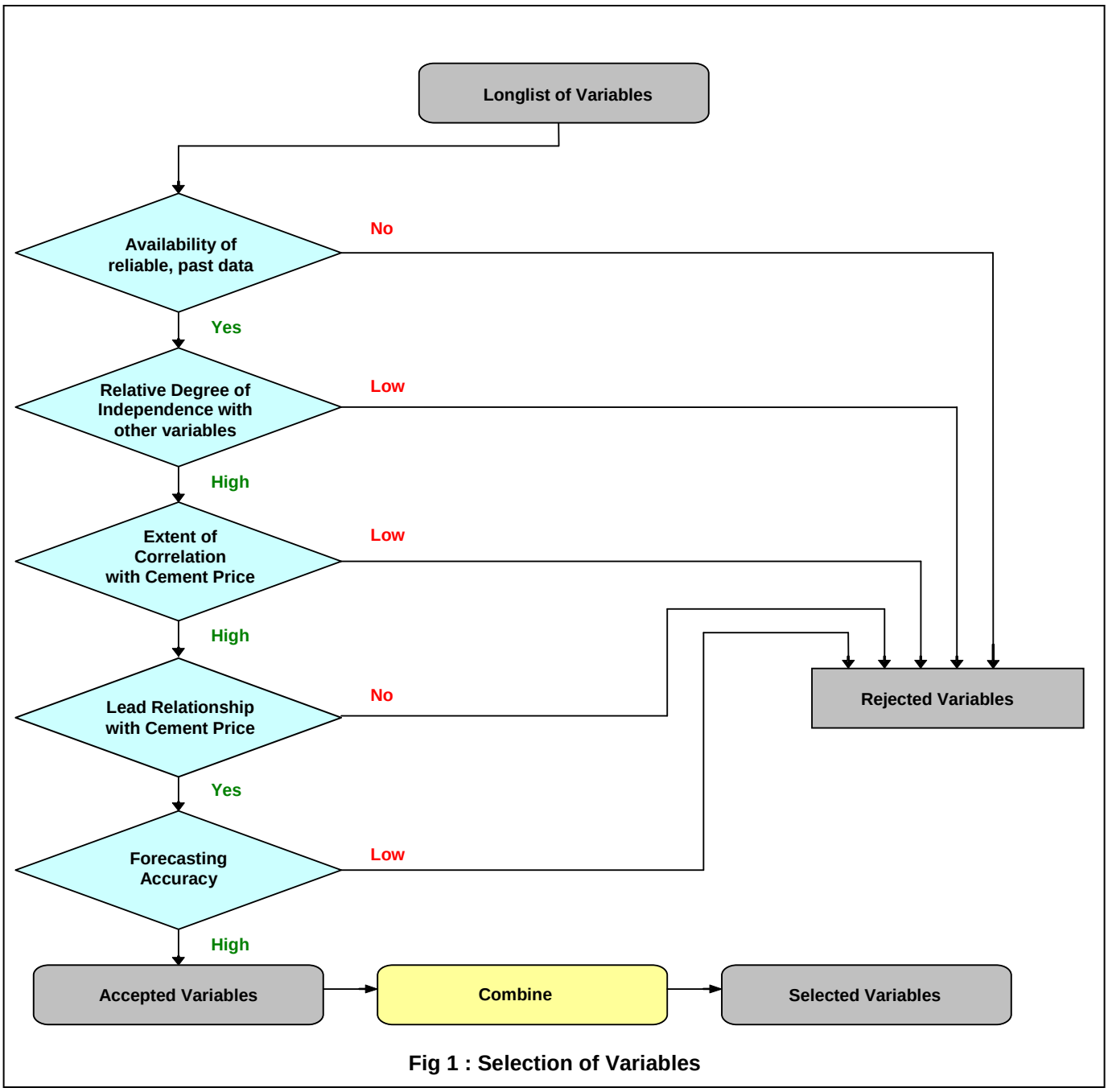

Fig 1 : Selection of Variables

Based on the above exercise, the 4 variables (basic & combined) finally chosen for forecasting the retail price of cement are shown below :

- Domestic Demand-Supply Gap
- Industry Consolidation in India
- Cement Consumption of Individual House Builders as a % of Total Cement Consumption in the Indian Cement Market.
- Inertia of Cement Retail Price

A discussion on each of these is presented hereafter.

1) **Domestic Demand-Supply Gap**

This variable was formed by taking the difference between two basic variables viz. Domestic Demand and Domestic Supply. Excess demand, in isolation, generally creates an upward pressure on the Cement Retail Price.

Domestic Demand : Past figures of cement consumption were assumed to have been equivalent to demand. In order to forecast domestic cement demand, a total of 6 lead variables, as measured over a period of 10 years, were employed to determine the ones most suited for prediction. A total of 26 curve fits, employing different subsets of these variables, were also experimented with. Theil's Coefficient was used as the accuracy measure for checking forecasts against actual results. The results showed that, in India, domestic demand was best predicted by a power curve (elasticity 1.2) linking it to real GDP. Based on GDP forecasts (selected from a variety of estimates made by The Economist, The Planning Commission, etc.) yearly estimates of domestic demand were derived.

Domestic Supply : Past figures of cement consumption were assumed to have been equal to domestic supply after adjusting these for exports and imports. The first step in deriving future supply was to estimate capacity additions and the timings at which these would appear. This was derived from Holtec's proprietary database on the Indian cement industry **Incem**. To get effective capacity in any year, the existing capacity was adjusted by adding the new capacity and subtracting estimates of capacity closures (that are expected to happen on account of technological obsolescence, material shortages or poor cost performance). Taking into account a variety of reasons such as input shortages, market constraints, transport bottlenecks etc, potential capacity utilisation figures were derived for each year. These were applied to the effective capacity values and subsequently adjusted for exports in order to arrive at yearly estimates of domestic supply.

Domestic Demand –Supply Gap : Our exercises show that in the period FY 2009 - FY 2015, while Domestic Demand would *most likely* to grow from 187 to 318 mio t, potential Domestic Supply would most likely grow from 195 to 335 mio t. The year-wise, potential Excess Supply is shown in **Figure 2**.

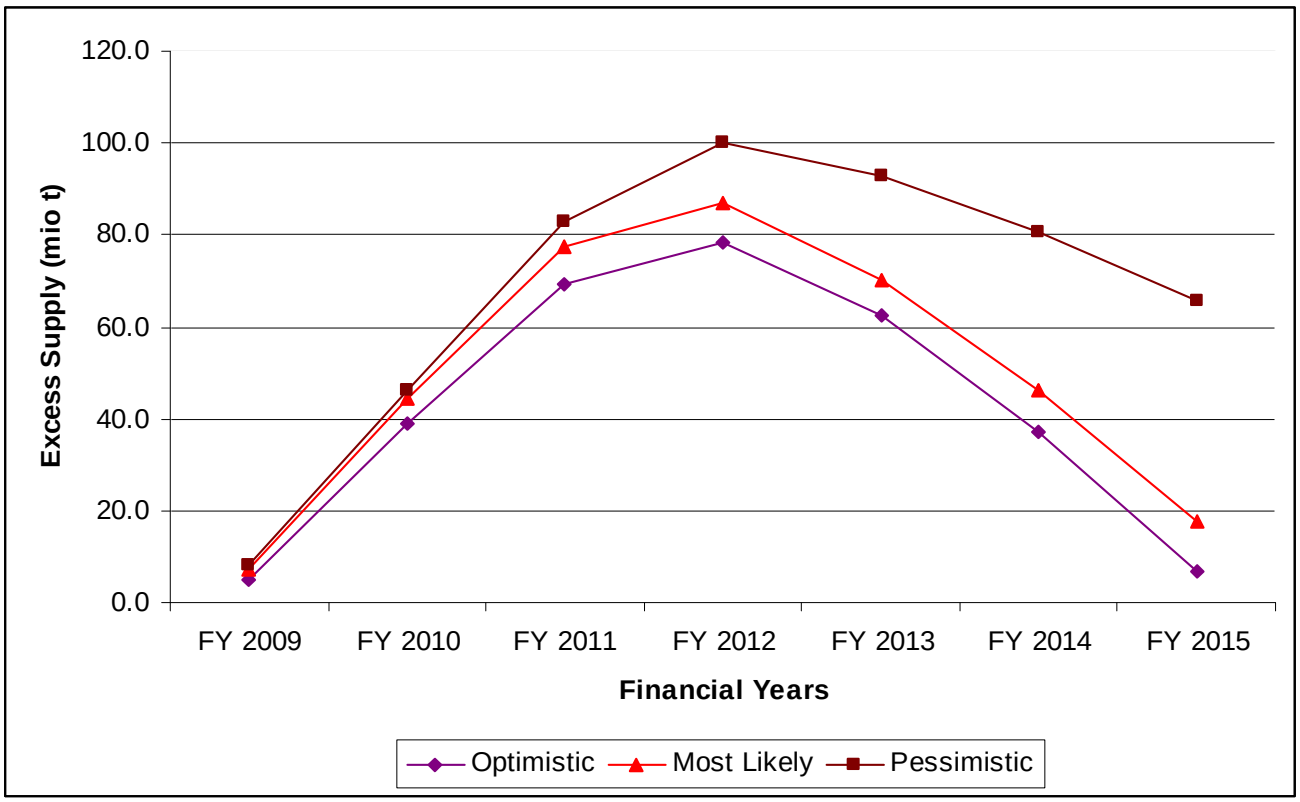

Figure 2 : Domestic Demand-Supply Gap

2) **Industry Consolidation in India**

It is believed that, in isolation, a higher degree of industry consolidation generally creates an upward pressure on the Cement Retail Price.

To measure the degree of consolidation in the Indian cement industry, the Herfindahl Index (H) was used. This metric enjoys a high degree of application in U.K. as well as U.S.A. where its absolute value, or change thereof, spurs government intervention targeted towards dispersing cartels. In the wake of recent pronouncements made by the Government of India, alleging cartelization in the Indian cement industry, this metric was preferred over others. By definition,

$$H = \sum_{i=1}^{n} s_i^2$$

where "s_i" is the market share of firm "i" in the market, and n denotes the number of firms. An absolute monopoly would have a score of 1.00 (or 100%), whereas perfect competition, with infinite players, would have a score of 0.00 (or 0%). For past data, actual market share data was used after considering all mergers & alliances. For the future, projected capacities were considered. Contrary to popular perception, the Index is seen to actually dip over the next few years, which seems to indicate that **in the times ahead, the industry would be moving towards de-consolidation or higher fragmentation**. The reason is that many new players are in the process of implementing large capacity projects.

Figure 3 depicts the behaviour of the Herfindahl Index as seen in the past and is expected to develop over the period FY 2009 - FY 2009.

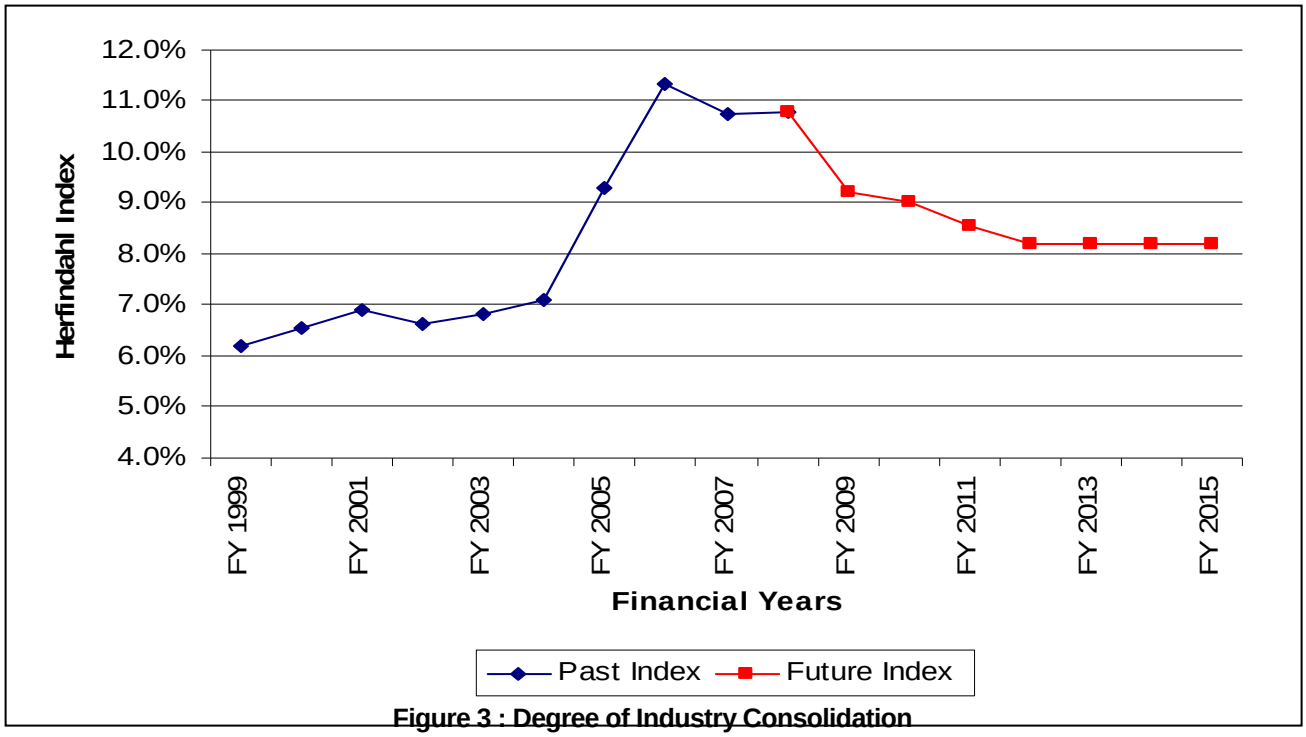

Figure 3 : Degree of Industry Consolidation

3) **Cement Consumption of Individual House Builders**

It has been observed in the past that as the Cement Consumption in the Individual House Building segment reduces, in relation to the Total Cement Consumption, an upward pressure is created on Cement Retail Price. It is

possible that on account of “affordability”, this segment is more sensitive to prices, and prefers postponing the timing for investment.

Figure 4 depicts the past and future percentage Cement Consumption of this segment in relation to the Total Cement consumption.

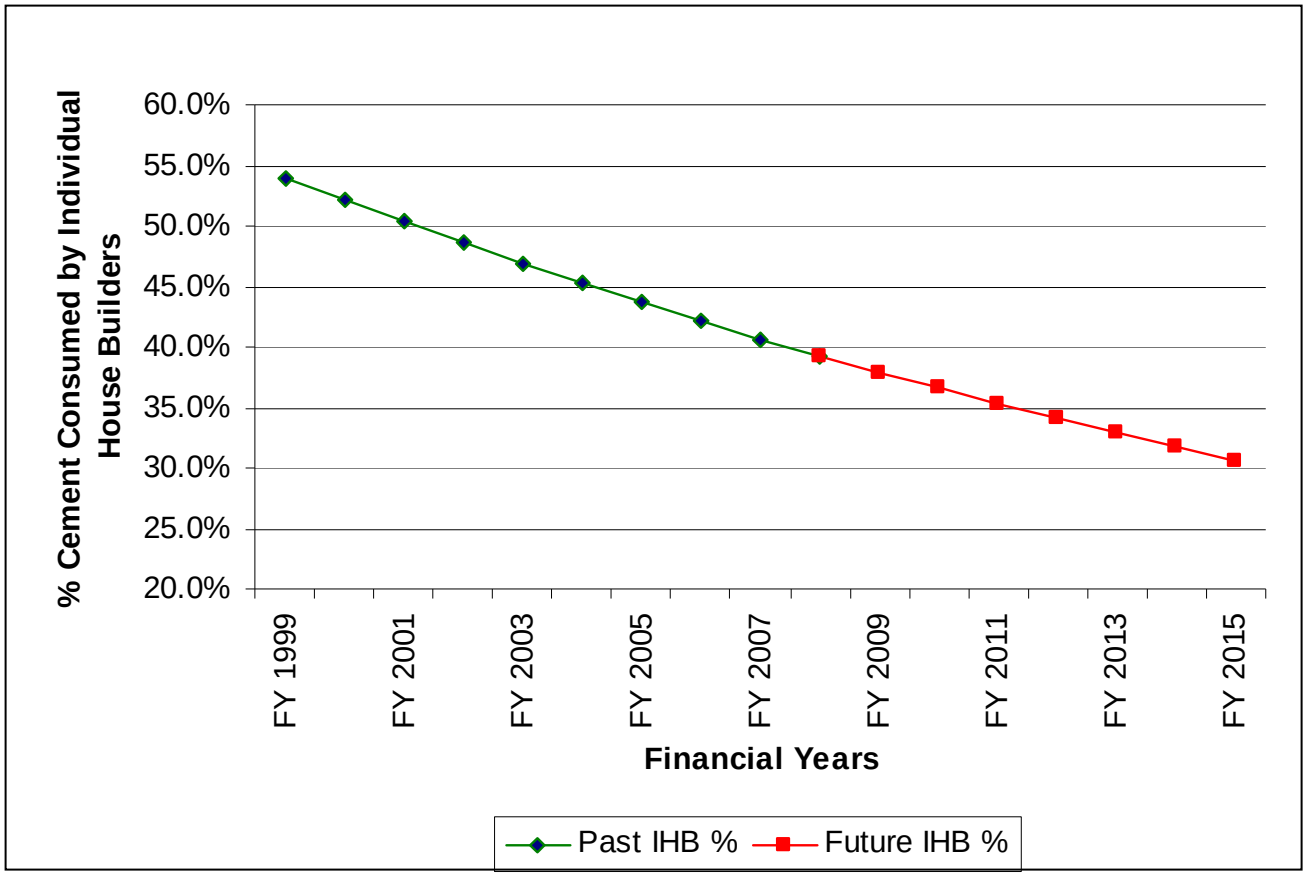

Figure 4 : Cement Consumption of Individual House Builders

4) **Inertia of Cement Retail Price**

In order to consider the possible anchoring effect that previous Cement Retail Prices have on future Cement Retail Prices, the relationship of the current period price P_t was tested against past prices, P_{t-1}, P_{t-2}, etc. Exponential smoothing was used as a technique to study cumulative relationships. The basic concept of exponential smoothing is that the value of a variable in a particular period is influenced by its values in previous periods with higher weights being given to proximate periods as compared to those more distant.

Various degrees of exponential smoothing are possible. In the case of the Indian Cement Industry, double exponential smoothing gave the minimum residual error with an alpha value ~ 1.00 and a gamma value = 0.91. The high alpha value indicated that the Cement Retail Price in a particular year was almost fully influenced by the Cement Retail Price in just the previous year with other years having almost no bearing. This implied that the memory for price did not run deep; only last year’s price exercised a significant influence. Thus, only P_{t-1} needed to be considered.

Figure 5 depicts the behaviour of the actual Cement Retail Price in the past, vis-à-vis that predicted by exponential smoothing.

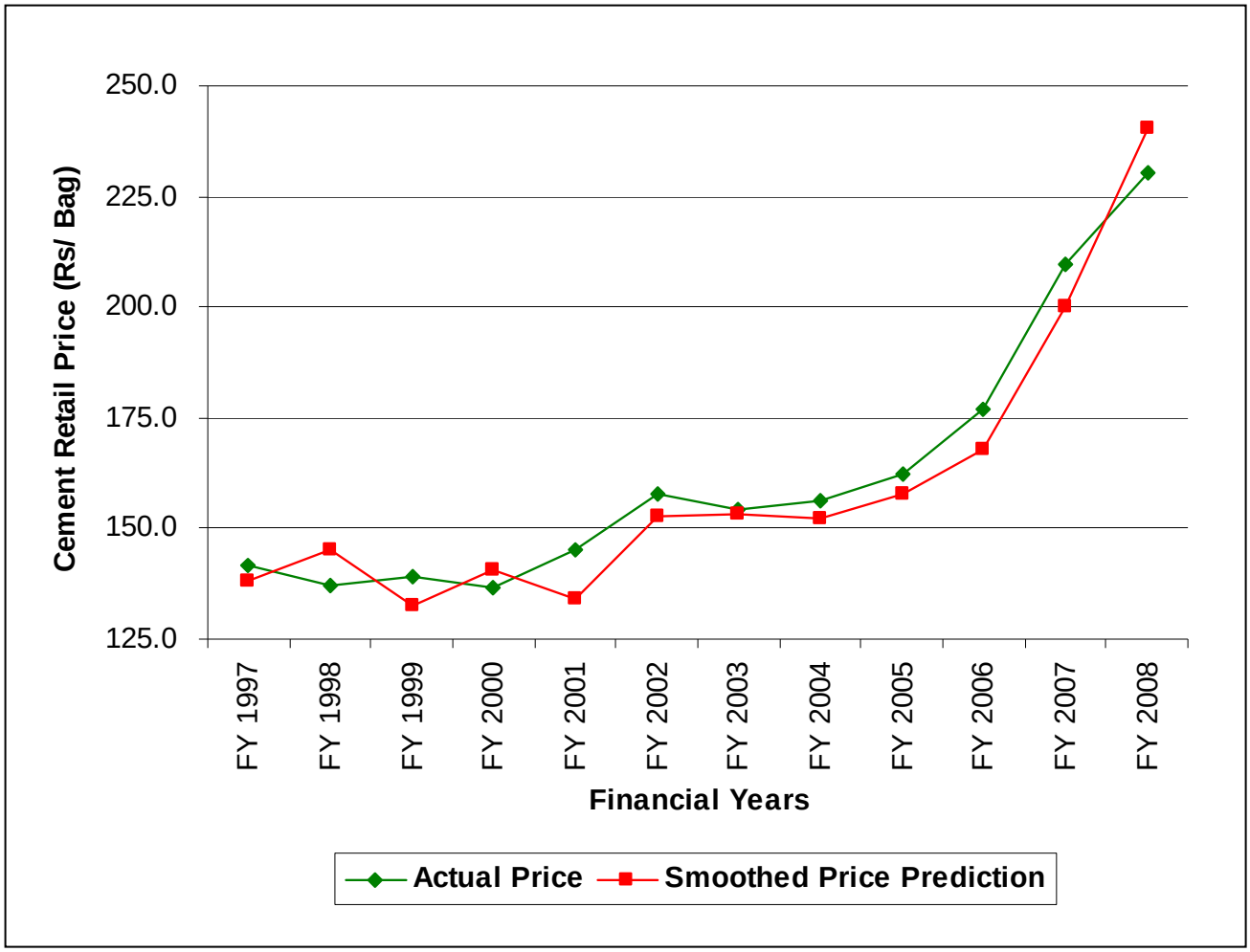

Figure 5 : Price Inertia & Exponential Smoothing Predictions

4. TERMINOLOGY USED IN FORECASTS

1) **Static & Dynamic Forecasts**

The word **Static** is only used to denote that the forecast for the dependent variable, viz. the Cement Retail Price, used single point, deterministic estimates of the independent variables. It was based on a Multiple Linear Regression Model, bounded by floor and ceiling values of the Cement Retail Price.

The word **Dynamic** is used in the context of stochastic distributions being assumed for the independent variables and simulation being used as a technique to determine outcomes based on the already derived Multiple Linear Regression Model.

2) **Optimistic and Pessimistic Occurrences**

The words **Optimistic** and **Pessimistic** in this paper refer to the specific view point of the cement industry. With the price of cement being the main concern of this paper, occurrences that are expected to result in a price increase are referred to as being “optimistic” with the converse being true as well. As an example, delays in project commissioning, in the context of this paper would be branded as an “optimistic” occurrence, since it would to selection.

3) **Real Cement Retail Prices**

The term Real refers to the time context. All forecasts made for the Cement Retail Price are at FY 2008 prices and do not consider the effect of normal inflation indices.

5. THE STATIC FORECAST

Based on the multiple linear regression technique the model for forecasting the Cement Retail Price is shown below

$$\mathbf{P} = 93.42 + 0.41\mathbf{G} + 2.9662\,\mathbf{H} - 0.8204\,\mathbf{I} + 0.56\,\mathbf{E}$$

where

P = Cement Retail Price in Rs/ Bag at 2008 prices

G = Demand-Supply Gap in mio t

H = Herfindahl Index as a percentage

I = Cement Consumption of Individual House Builders as a percentage of Total Cement Consumption

E = Exponentially Smoothed Price Forecast in Rs/ Bag at FY 2008 prices

The strengths of the independent variables in the multiple linear regression model are shown below :

Parameter	Intercept	G	H	I	E
Coefficient (Slope)	93.4	0.41	296.62	(82.04)	0.56
Standard Error	114.8	0.7	351.3	118.9	0.2
Confidence Level (%)	78	71	79	74	98

It is evident from the above that the strongest variable is the Exponentially Smoothed Price Forecast as this can be stated with a confidence level of 98%. This is followed by the other variables H, I and G all of which enjoy confidence levels between 70 - 80%.

The multiple linear regression model displays the following statistics :

Statistic	Value
R^2	95.2
Standard Error of y-estimate	8.22
Degrees of Freedom	7
F-value (observed)	34.34
ss resid	472.8
ss reg	9,278

While the confidence levels of the individual, independent variables are as stated earlier, when grouped together, the overall model becomes very strong with an F-test yielding a confidence level of 99.99% at a significance level of 0.10 (single tail).

Figure 6 shows the Cement Retail Price obtained by the static model. It additionally shows the values predicted by it for the past vis-à-vis those observed actually.

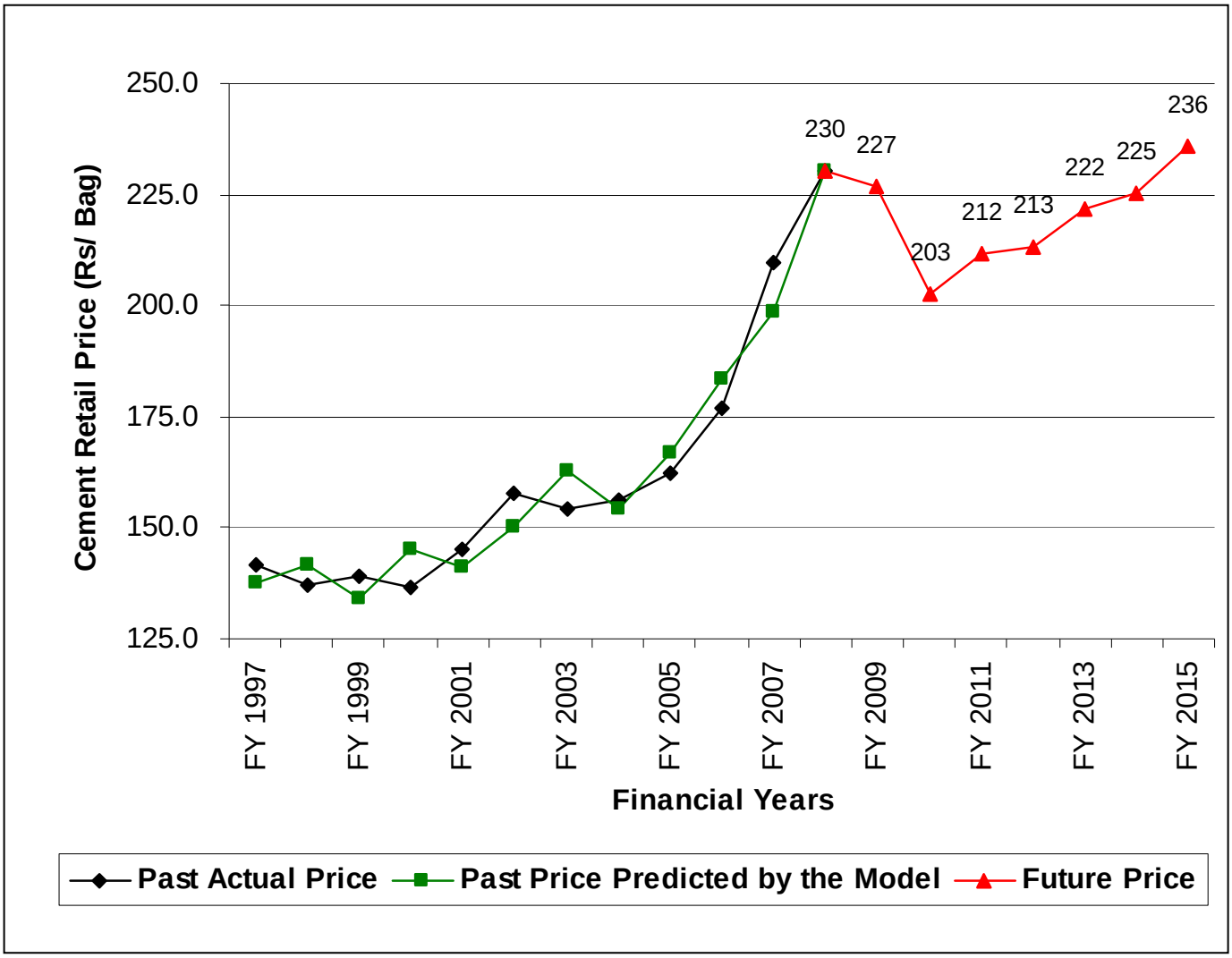

Figure 6 : Price Forecasts using the Static Model

6. THE DYNAMIC FORECAST

In the absence of confirmatory data, the distributions for the independent variables were assumed to be triangular, with associated pessimistic, optimistic and most likely values. Monte Carlo Simulation was used as the quantitative technique to forecast the Cement Retail Price. For the combined variable, the Demand-Supply Gap, interactions between the pessimistic, optimistic and most likely values for the Demand (using different values of GDP) as well as the Supply (by assuming advancement and deferment of capacity creation by 10 – 25%) were also subjected to Monte Carlo Simulation.

Thorough convergence was achieved after about 900,000 simulations. The results of the Dynamic Forecast are shown in **Figure 7**. Other than the mean of the simulated outcomes in each year, the figure also shows the 90% confidence level bands. These bands are indicative of the fact that 80% (i.e., 90% less 10%) of simulated outcomes lie between these upper and lower bands. It was observed that subsequent to FY 2009, the price floor was invoked. This is fairly likely on account of the fact that it is unreasonable to assume that producers would continue selling cement, on a sustainable basis, below production+ ownership costs. Had the price floor criteria not been invoked, the model predicts that the Cement Retail Price would have dropped to a level of Rs 105/ bag in FY

2012. Thereafter, there would have been a gradual recovery leading to a value of about Rs 164/ bag in 2015.

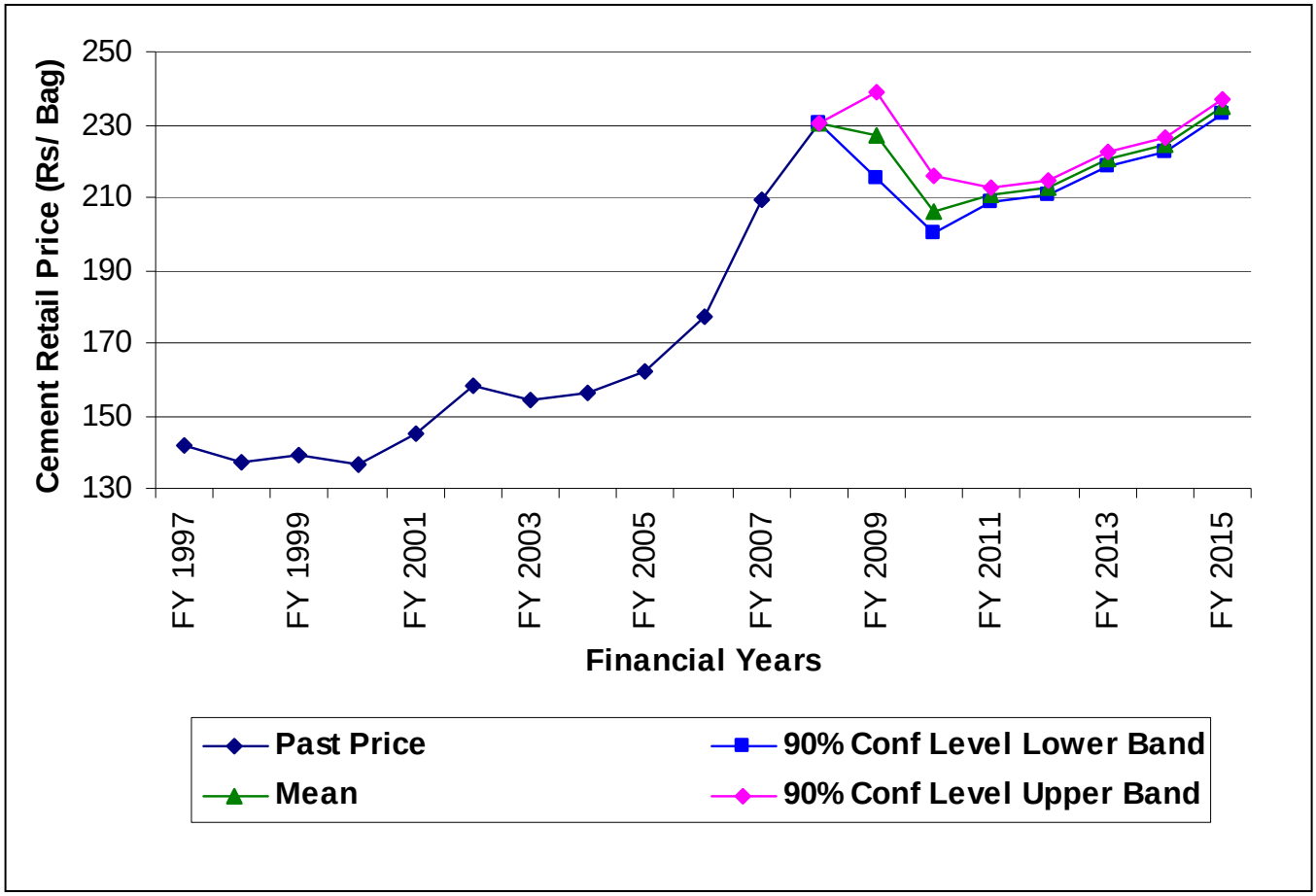

Figure 7 : Price Forecasts using the Dynamic Model

The simulation exercise also made it possible to study the probability distribution of the Cement Retail Price at the two terminal points (viz. FY 2009 and FY 2015) and one intermediate point (Say, 2011) of the forecasting period. The results are shown in **Figures 8, 9 and 10**, respectively

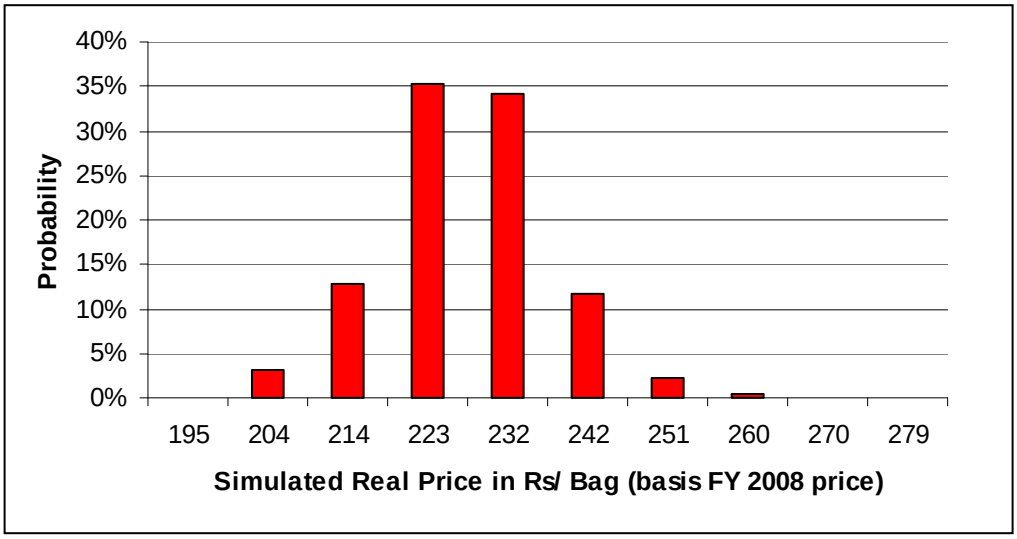

Figure 8 : Probability Distribution of Cement Retail Price in FY 2009

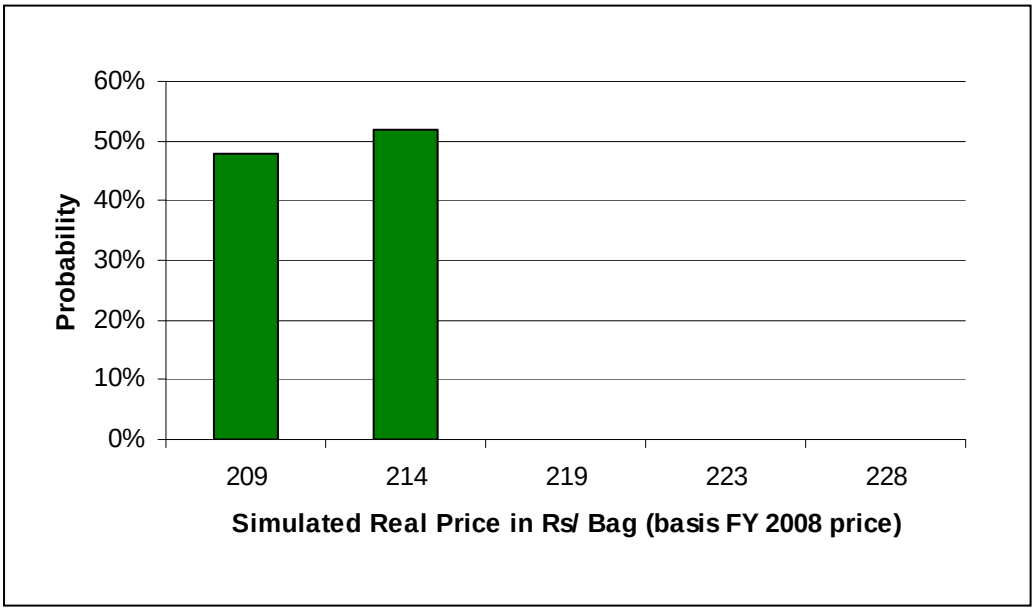

Figure 9 : Probability Distribution of Cement Retail Price in FY 2011

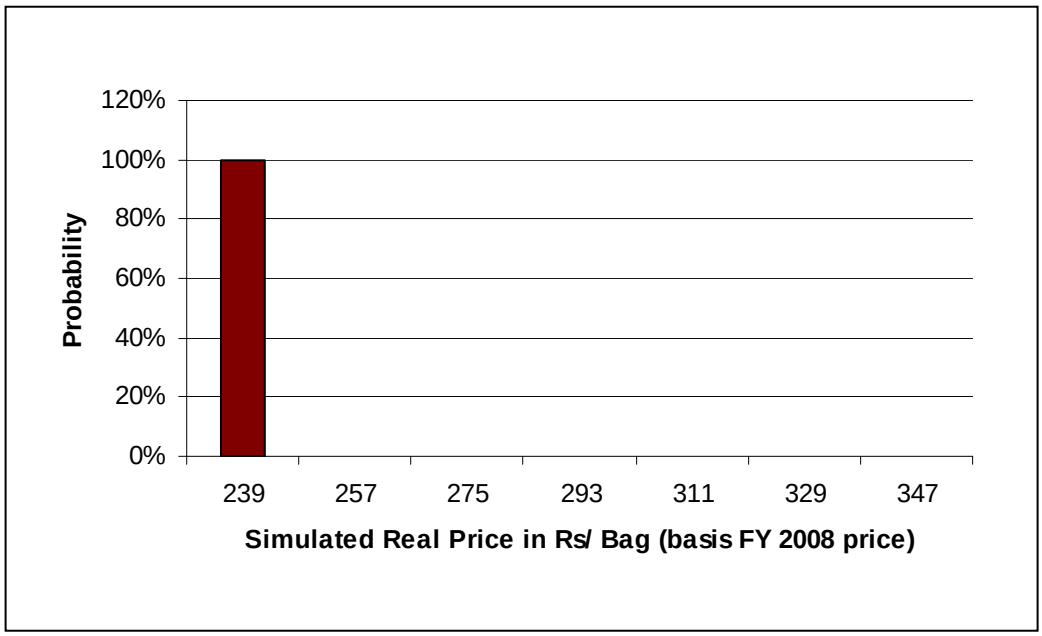

Figure 10 : Probability Distribution of Cement Retail Price in FY 2015

At first glance, the information shown in Figures 8, 9, 10 appear strange. One would normally expect that the further we are at a point of time the higher would be the uncertainty, and therefore the spread of the probability distribution. The real reason is that the left tails of the respective distributions get truncated on account of the imposition of the floor price which is driven by production & ownership cost considerations.

MODEL RESULTS

The results of the application of the static and dynamic models are shown in the following table:

Financial Years	Cement Retail Price in Rs/ Bag (basis FY 2008 Price)			
	Static Model	Dynamic Model		
		Lower 90% Confidence Level	Mean	Upper 90% Confidence Level
2009	227	215	227	239
2010	203	200	206	216
2011	212	209	211	213
2012	213	210	213	215
2013	222	219	221	223
2014	225	222	224	226
2015	236	233	235	237

It may be observed that the outcomes of the Static Model and the Dynamic Model are not drastically different.

8. CONCLUSIONS

The methodology adopted in this paper attempts to give an insight of the development of real (base - FY 2008 prices) Cement Retail Prices in the period between the current Financial Year, FY 2009 and FY 2015 at the national level. The outcomes presented are reflective of the average price in a year and not the cyclic highs and lows.

The steep climb from Rs 170 to Rs 230 per 50 kg bag, witnessed in the period FY 2005 to 2008, is likely to be arrested with prices falling to Rs 227 $\pm$ 12 in FY 2009. As capacity additions begin to overtake the increases in domestic consumption, a further deceleration is likely with prices falling to Rs. 206 $\pm$ 6 in FY 2010. In the opinion of the authors, the prices are not expected to fall further, despite significant capacity surpluses, because the model kicks in a cost-based floor, derived from a sum of production & ownership cost inputs. Such a hypothesis appears justified, not only on account of the cost increases being witnessed in scarce resources (e.g. fuel), but also on account of the high capex being expended in creating capacity on which a reasonable return is naturally expected. With the envisaged capacity creation being much in excess of the existing capacity, producers are unlikely to allow a free fall in price.

The basic methodology, being fairly broad-based, can find application in regional pockets within India as also in other countries. The selection of variables, in such environments would however differ.