Dear LDS,

In order to test the hypothesis of whether churn is driven by the customers' price sensitivity, we would need to model churn probabilities of customers, and derive the effect of prices on churn rates. We would need the following data to be able to build the following models.

Data needed:

1. Customer data - which should include characteristics of each client, for example, industry, historical electricity consumption, date joined as customer.

2. Churn data - which should indicate if customer has churned.

3. Historical price data which should indicate the prices the client charges to each customer for both electricity and gas at granular time intervals.

Once we have the data, we would need to engineer features based on the data that we obtain, and build a binary classification model (e.g. Logistic Regression, Random Forest, Gradient Boosted Machines to name a few), picking the most appropriate model based on the tradeoff between the complexity, the explanability, and the accuracy of the models. Based on the model picked, we would be able to understand the direction and magnitude of the impact of prices on churn rates, as well as the relative importance of prices compared to other factors. Furthermore, the model would allow us to size the business impact of the client's proposed discounting strategy.

Yours respectfully,

JAFSIA Elisee