Appendix 10.7: Domestic pricing analysis

Introduction

1. This analysis adopts a regression approach to compare domestic energy prices (bills) across suppliers, breaking down the observed bill differences into a set of price drivers. The objective is to assess the ranking of suppliers and the scale and significance of differences between suppliers, while controlling for differences in customer mix. We also assess differences by region and payment method.

2. This analysis controls for:

   (a) consumption levels (some suppliers may have a customer base with higher consumption and therefore lower unit revenues);

   (b) regional customer mix (some suppliers may have a larger proportion of their customer base in regions where network charges are higher);

   (c) payment method mix; and

   (d) proportion of customers subscribing to long-term tariffs (tariffs with a fix of over two years, which tend to come at a premium).

3. Comparisons of average unit revenues across suppliers cannot control for all of the above factors because sufficiently disaggregated data is not available.

Data

4. The data used is tariff data from the analysis of the potential gains from switching. Appendix 7.4: Analysis of the potential gains from switching has a detailed description of the dataset. In short, this data is a list of all tariffs customers of the Six Large Energy Firms were subscribing to at end-of-quarter snapshots between Q1 2012 and Q2 2014; and the bills of these customers calculated at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles and the mean of the consumption distribution within that supplier, region, year and payment and tariff type.

5. Each tariff has multiple entries where each entry represents a different price. Price can differ within a tariff because some customers receive a specific

discount while others do not. For each of these prices we know the total number of customers subscribed.[1]

6. The data includes all Six Large Energy Firms, and the white label tariffs of Centrica and SSE. The white labels are pooled together with the 'parent' firms in the analysis.

7. Only the 25th, 50th and 75th consumption percentiles are used in this analysis, but the lower and higher percentiles are added back for one of the sensitivities. This analysis only covers dual fuel customers.

Methodology

8. The approach is known as an analysis of variance. It decomposes the bill into various drivers. We implement this by means of a regression, where the dependent variable is an annual bill. In our main regression the explanatory variables are:

   (a) the identity of the supplier;

   (b) the level of annual electricity and gas consumption in kWh;

   (c) the payment method (direct debit, credit and pre-payment);

   (d) whether the tariff is Economy 7; and

   (e) the region.

9. The regression is estimated using ordinary least squares, weighted by numbers of accounts[2] and standard errors are clustered by tariff. All explanatory variables except for *(b)* are dummy variables (binary indicators). Our main regression is a pooled regression including data from each quarter from Q1 2012 to Q2 2014. The computed differences between suppliers, payment methods and regions are therefore an average across quarters and may reflect changes over time in the composition of variables which are not controlled for. We test for significant differences between the supplier and payment method variables.

10. To ensure robustness of the results, we conducted sensitivity analysis using a number of alternate specifications and time periods. This consists of the following modifications (results are not reported):

---

[1] See Appendix 7.4: Analysis from the potential gains from switching, Annex C: Consumption assumptions, for the methodology of allocating customers across the different levels of consumption.
[2] Therefore, for example, tariffs with more suppliers and tariffs will get more weight in the regression.

(a) Using all the data, but additionally adding a control variable for long-term, fixed-term tariffs. We express this as a dummy variable indicating whether the contract length is more than 24 months.

(b) Estimating the regression for each quarter separately.

(c) Adding dummy variables for each quarter to the main regression.

(d) Adding interactions of quarterly dummy variables with supplier indicators.

(e) Adding a dummy variable for online tariffs.

(f) Adding a dummy variable for whether the supplier is the regional electricity incumbent.

(g) Estimate the regression for standard variable tariffs (SVTs) only. This allows us to compare the suppliers' SVT prices alone.

11. We note that compared to the average unit revenue analysis, this approach relies on the tariff dataset, which:

(a) does not include all tariffs (excludes complex time of use tariffs, green tariffs and erroneous data);

(b) accounts for the majority, but not all discounts (for example, prompt-pay discounts are not included in the data); and

(c) presents the issue that we do not have the actual consumption levels for each tariff; instead the consumption levels of tariff families (as defined by supplier, region, payment method, tariff type and year) are assigned to each tariff.

Results

12. [✂]

13. The key results are:

(a) [✂]

(b) [✂]

(c) [✂]

(d) [✂]

(e) [✂]

Table 1: Regression results

[✂]

Source: CMA analysis of supplier data.
Notes:
1. [✂]
2. [✂]

Figure 1: Differences across regions controlling for supplier, consumption, payment method, and economy 7 status

[✂]

Source: CMA analysis of supplier data.
Notes:
1. [✂]
2. [✂]