

CC Market Investigation – Opinion on SSNIP test methodology

1 I am Ron Smith, Professor of Applied Economics at Birkbeck College, University of London. I teach postgraduate statistics and econometrics at Birkbeck and have published extensively in statistics and applied econometrics on a variety of topics in Journals such as the Economic Journal, American Economic Review and the Journal of the American Statistical Association. I am also an Associate with Frontier Economics and act as an adviser in relation to empirical matters. In particular, I have provided Frontier with some empirical advice on matters arising from the current Groceries Inquiry including the work on the SSNIP test.

2 Freshfields Bruckhaus Deringer recently asked me to provide an expert opinion on the SSNIP test methodology developed by Frontier and Tesco in relation to geographic market definition in the Competition Commission's market investigation into the UK groceries sector. In doing this I have reviewed: (1) a submission entitled "Further Evidence Supporting our View That There is a National Geographic Market"; (2) a presentation by Decision Technology "Total Customer Cost: Superpanel Models"; and (3) a submission entitled "Response to Emerging Thinking, Market Definition Working Paper". I have been asked to give my professional opinion on the conceptual arguments and statistical implementation contained within them. My opinions are based on these documents; I have not looked at the data or replicated the results.

3 I have set out my views on five aspects of Tesco's work:

- a. Tesco's methodology;
- b. The information that Tesco relies upon;
- c. The assumptions that underpin Tesco's work;
- d. The questions raised by the CC; and
- e. The results and cross-checks employed by Tesco.

Tesco's methodology

4 The purpose of document (1) is to identify the appropriate size of the geographic market(s) for the purpose of the grocery inquiry. The document rightly states that the correct test to define a market is to ask whether a hypothetical monopolist (HM) could sustain a small but significant and non-transitory increase in price (SSNIP).

5 The most difficult part of conducting a SSNIP test formally is to determine the extent to which customers would be willing to switch from one geographic area (or product) to another in response to a SSNIP, that is, estimating the price elasticity of demand. The difficulties associated with this include the identification problem, sensitivity to omitted variables, and the choice of functional form. All of these problems can lead to estimates of own and cross-price elasticities being of incorrect sign or implausible magnitude.¹

¹ Such anomalies are discussed by A. Bonfrer, E.R. Berndt and A. Silk, "Anomalies in estimates of cross-price elasticities for marketing mix models: theory and empirical test" NBER working paper 12756, 2006.

6 Moreover, estimating the price elasticity of demand particularly difficult if there is a lack of geographic variation in prices. If there is no variation in prices geographically, one will not observe how customers react to changes in the relative price of stores in one area compared to stores in another and it will be impossible to estimate the price elasticity of demand.

7 Tesco avoids these difficulties by using a model of customer choice which is driven by the relative costs of using different stores. In summary, Tesco explain that a customer will switch from his local store to a more distant store following a SSNIP if the extra cost of travelling to the more distant store is less than the extra cost associated with purchasing (the now more expensive) basket from his local store. This is a sensible simplified model of customer behaviour.

The information that Tesco relies upon

8 The information that Tesco uses in its modelling is set out clearly in document (1).

9 There are three main sources of information that Tesco has relied upon in its modelling.

- a. A database of store and customer location information.
- b. TNS Superpanel data.
- c. Margin information.

10 These sources appear to contain right type of information for conducting this type of modelling exercise. I have not interrogated the sources myself, but I have been told that Tesco has submitted them to the Competition Commission and so it would be easily verifiable if necessary.

- a. A central component of this analysis is the distance between customers and stores. Tesco has analysed customer switching at the Census Output Area level (COAs). I have been told that COAs are highly disaggregated population centres (around 100 customers in each) and this would allow for at least as great a level of precision compared to alternative population centres (such as postcode areas).
- b. TNS Superpanel data is used to calculate a distribution of basket sizes and is also used in the econometric work set out in document (2). The dataset includes information about the actual shopping baskets of over 10,000 people across a large number of different stores. This is a sensible set of data to evaluate customer behaviour and I understand that it is used regularly by Tesco in its decision making.
- c. The revenue and cost information that Tesco has used appears appropriate for estimating a margin suitable for the SSNIP test.

The assumptions that underpin Tesco's work

- 11 It has been necessary for Tesco to make a number of simplifying assumptions to operationalise the SSNIP test. This is normal in applied work. The assumptions that Tesco uses in its modelling are, in my opinion, reasonable. Collectively, they do not appear systematically to bias the subsequent calculations towards making the SSNIP less profitable and thus do not appear to be overly supportive of Tesco's position.
- 12 The assumptions that Tesco uses in its modelling are also clearly set out in documents (1) and (3).
- a. *The hypothesised local market uses the CC's previous market definition for "one stop shopping".* This statement merely sets out Tesco's starting point for the test and is not material in the calculations that follow.
 - b. *The analysis explores "one stop shopping" only and uses the CC's effective competitor set.* This implies that customers regard all stores in the competitor set as equal substitutes to one another, that is, it is only distance that distinguishes them from one another. Of course, other features of stores also matter to consumers, but this is a reasonable starting point for the analysis.
 - c. *All customers shop at their nearest store.* Although customers may not shop at their nearest store, this assumption is consistent with the idea that travel times are an important delimiter of geographic markets. It is not clear that this assumption biases the results in favour of making the SSNIP less profitable – the decision to switch depends upon the time between a customer's nearest store inside the market and the nearest store outside the market.
 - d. *All stores in the hypothesised market raise their price by 5%.* It is typical to assume a SSNIP of 5% and it is correct to assume that every store in the competitor set "belongs" to the HM. In principle, it could be possible for the HM to increase the price of stores in its "chain" by differing amounts – for example, by charging higher prices to customers that have further to travel to a store outside of the candidate market. This would increase the profitability of a SSNIP and so lead to estimated market sizes being too wide, although the magnitude of this effect is unclear. However, determining the appropriate values of the price increase across many stores would be a highly complicated exercise. This assumption is a pragmatic step.
 - e. *Customers switch in response to a price increase if the extra transport costs from switching are lower than the price increase.* As noted above, this is a sensible model of consumer behaviour and is consistent with the idea that travel times are an important delimiter of geographic markets.

Document (2) describes a model of customer choice, whereby customers select a store on the basis of its “total cost”. Specifically, Decision Technology assumes that a customer prefers to shop at a store which minimises the cost that he incurs given the shopping mission that he embarks upon. This is a standard description of rational customer behaviour.

To apply this framework empirically, Decision Technology estimate a conditional logit model which explains where a particular basket was purchased as a function of drive time, store size, fascia, and mission size. Conditional logit models are often used to analyse the drivers of choice when customers can choose from a number of alternatives. It is an appropriate approach to answer the question: what is the effect of travel time on store choice?

Decision Technology undertook a number of statistical robustness tests. They compared Akaike Information Criterion scores from a number of models and also compared the predicted market shares against actual market shares. In addition to these tests one could also check that the estimated coefficient for drive time is statistically significant and has a similar magnitude across different model specifications (for example, by including or excluding particular variables).

In document (1) Tesco assumes that all customers place a value of [1] on their time, this is value implied by the work conducted by Decision Technology. I have not replicated its analysis, however I understand that Tesco found an error in the application of the methodology, which when corrected gives a revised estimate of [1]. This error does not affect my views on the overall robustness of the methodology. In document (3) Tesco assume a cost of travel of the revised figure of [1].

The analysis uses a single estimate of the value of time across all customers. This is a simplification of reality. It is likely that the value of time varies from customer to customer. However, this simplification is not, in itself, likely to bias the results of the subsequent calculations towards making the SSNIP less profitable. This would only occur if the value of time increases significantly with basket size or with the switching opportunities available. It would be possible to test this empirically by including an interaction term between drive time and basket size in the conditional logit model. In Annex B of document (3) Tesco investigate this possibility and find that the cost of travel is inversely related to basket size.² This would imply a conservative bias in the Tesco results.

There is a question as to whether this is reasonable. My intuition is that it may be customers who purchase small baskets when they are short of

² This is discussed in paragraph 38 and the related footnote 64 of Annex B.

time – such as convenience shoppers – who have a higher cost of travel time. This would be consistent with Tesco’s empirical findings.

- f. Customers switch their entire basket.* This implies that customers would not substitute part of their basket to alternative retailers in response to a SSNIP. As in (b), this appears to be a reasonable assumption to operationalise the test. I cannot think of a sensible reason why a customer would switch only a proportion of their basket rather than the whole basket, given that the analysis explores one stop shop stores only.
- g. Tesco does not consider the effect of purchases which are not one stop shops.* Although non-one-stop customers would face price increases if it were impossible for the HM to distinguish between them and one-stop customers. It is not clear that this will bias the SSNIP test in any particular direction. On the one hand, as Tesco notes, non-one-stop customers may switch to stores which are not one stop shops within the candidate market and make the SSNIP less profitable. On the other hand, these customers may have smaller baskets and so the effect of a 5% price increase will not necessarily encourage them to switch and make the SSNIP more profitable.
- h. There is variation across customers.* Tesco has applied a common distribution of baskets to all COAs. In reality, the actual distribution of baskets is likely to vary throughout the country. Again, it is not clear that this will bias the SSNIP test in any particular direction. In some areas more baskets would switch in reality, in other areas, less.
- i. Some costs are reduced when customers switch.* Tesco correctly identifies that the relevant margin to use in the SSNIP test depends upon the extent to which it would be possible for the HM to reduce its costs by serving fewer customers. The costs identified, and the allocation to variable, semi-variable and fixed cost categories, appear reasonable (although no underlying evidence was provided in document (1) as to why the allocation proportions that have been used were correct).
- j. In document (1) Tesco use a margin figure of [1]%, whereas in document (3) Tesco use a revised margin of [1]%, which they state more accurately reflects the extent to which they view certain costs are variable. I cannot comment on whether these figures are of the right order of magnitude or whether they are robust, but the methodology employed seems sound.*

It is possible that margins could vary throughout the country. It is not clear that this assumption would materially bias the results in any particular direction.

The questions raised by the CC

- 13 I have been told that the Competition Commission questioned the omission of an “outside option” from Decision Technology’s work during a staff meeting held on 31st January 2007. The purpose of including an outside option would be to allow for the possibility that a customer could choose to purchase a basket at a retailer not included in the dataset or to not purchase a basket at all. I have been asked to comment on the effect of omitting an outside option on the cost of time estimate used in Tesco’s analysis.
- 14 In relation to the econometric analysis carried out by Decision Technology, the appropriate outside option would be the possibility for customers either not to shop for groceries at all, or to shop at stores outside the set monitored in the Superpanel (which I understand would be primarily small stores such as convenience stores).
- Outside options are normally considered in situations where there is a clear alternative to the behaviour captured by the data. It is not clear that an outside option – in the form of not shopping – is relevant to the situation of grocery shopping since (i) customers shop regularly – it is not clear how one could ever identify a decision “not to shop” – and (ii) the TNS Superpanel data covers the vast majority of the stores that customers ever use for grocery shopping.
 - It is not clear that – even if customers did use an outside option – this would lead to any bias in the coefficient on the cost of travel time. For this to be the case, it must be that those customers who chose to use the outside option would have costs of travel time that are either systematically greater or systematically smaller than average. I am not aware of reasons why this might be the case.
 - It seems unlikely to me that this outside option could realistically ever be taken into account given the data limitations. The TNS Superpanel data only includes the actual shopping trips of customers, not their decisions whether to shop or not. I am unaware of a dataset which contains information on the latter issue for grocery shopping.
- 15 In relation to the SSNIP test modelling itself, the exclusion of an outside option reduces the profitability of a SSNIP and hence lead to wider geographic markets than Tesco has estimated. This is because customers would be able to switch to smaller stores in the event of a SSNIP and those stores are not captured in the model.
- 16 In addition to this point, I understand that the CC has raised a number of further questions in writing (in a letter of the 15th February) relating to the assumptions underlying Tesco’s analysis. Tesco addresses these questions in Annex B of document (3). I have reviewed these questions and Tesco’s responses, and I summarise my conclusions below.
- a. *Possibilities which would be expected to result in a conservative bias.*
In a number of cases, the effect described by the CC would serve to understate the profitability of a SSNIP. These are the assumptions

relating to cost of travel, the exclusion of customers outside the hypothetical market, and the exclusion of switching to smaller stores.

- b. *Possibilities which could, as a matter of theory, either over- or underestimate the scope of the geographic market.* In a number of cases, the effect the CC describes would tend to reduce the overall precision of the SSNIP test (i.e. the accuracy of the estimation of the geographic market) because some factors which affect customers' switching decisions are not accounted for as a result of a simplifying assumption. These are the assumptions relating to store selection on the basis of factors other than distance (e.g. price or quality) and the restriction to "one stop shopping" baskets. These factors would not bias the results of the SSNIP test unless they were correlated with the number of stores available to customers who might switch. I see no intuitive reason why such a correlation would exist and do not therefore expect any systematic bias to the results from these factors.
- c. *Possibilities which would be expected to result in a generous bias.* The CC raises two possibilities which could be expected to bias Tesco's results in favour of wider geographic markets. The first is the possibility that the HM may be able to price discriminate between stores. If this were possible, the HM would be able to set prices to minimise customer switching and therefore increase the likelihood that a SSNIP would be profitable. The magnitude of any such effect is unknown. However, empirically, my view is that it would be difficult to determine the extent to which a price discriminating HM could profitably impose some vector of price increases without carrying out simulation of a large number of price scenarios in each hypothetical market, which is a complex exercise and may not be possible.

The second is the possibility that information costs would make customer switching less attractive. I agree with the CC and Tesco that this would tend to favour wider geographic markets. However, given the frequency of customers' grocery purchases, it seems reasonable to assume that information costs are not large enough to have a substantial impact on the conclusions of the analysis (given the robustness of the model conclusions to other sensitivity analyses as described below).

- 17 Overall, given that there a potential biases in both direction – and given the results of the sensitivity analysis described below – it does not seem to me that these issues are sufficient to call into question the overall findings of the analysis.

The results and cross-checks performed by Tesco

- 18 The results in document (1) show that, using the CC's definitions of local markets, a SSNIP test is unprofitable in 20 areas of the country. These results are based on a cost of travel assumption of [] and a margin assumption of []%.

19 Tesco provides two appropriate illustrations of the robustness of the model results in document (1).

- a. Firstly, Tesco shows the estimated propensity of customers to switch outside of the candidate market on a map. As is to be expected, customers are found to be more likely to switch the closer they are to the boundary of the candidate market and to alternatives outside of the market. This is a useful check to confirm that the model gives intuitively plausible results.
- b. Secondly, Tesco calculates the value that customers would have to place on their time in order to equate the HM gains and losses – that is, the value beyond which the SSNIP would be profitable and therefore the candidate market of the appropriate size. The analysis shows that in many of the 20 areas the value of time would have to be significantly above [] per hour. This provides comfort that the conclusions of the work are robust to changes in the cost of travel time variable.

20 In document (1), Tesco does not establish what the appropriate size of the geographic market should be by using the test. The test would be more informative if Tesco continued to widen the isochrones until the SSNIP became profitable for the HM. This is the analysis carried out for the national results in document (3), at least up to 30 minutes drivetime.

21 The results in document (3) show that, using the CC's definitions of local markets, a SSNIP test is unprofitable in 98% of stores in urban areas and 91% of stores in rural areas. These results are based on a cost of travel assumption of [] and a margin assumption of []%. Tesco presents SSNIP test results for hypothetical markets up to 30 minutes wide around all GB stores. The results show that a SSNIP test is unprofitable at 30 minutes in 89% of urban areas and 77% of rural areas. In these cases Tesco does not establish what the final size of the geographic market should be. These additional results confirm the main conclusion that markets defined as 10 and 15 minutes are too narrow in the majority of cases, although also points out that for a small number of stores the 10 or 15 minute drivetimes previously proposed by the CC would be appropriate.

22 I note that Tesco also present results on the previous assumptions of £[] and []% and that this does not materially affect their results. In response to a suggestion by the CC, Tesco also calculate their results on the basis of a lower []% margin assumption. Tesco are correct to state that using a higher margin assumption would tend to favour a profitable SSNIP, and would therefore make the presented results conservative. On the basis of the lower []% margin assumption, the SSNIP test is unprofitable 97% of urban areas and 90% of rural areas. This is helpful evidence that Tesco's conclusions are robust to small changes in the margin assumption.